



# Caractérisation structurale et fonctionnelle d'une lectine de type-C des cellules de Langerhans : La Langérine

Eric Chabrol

## ► To cite this version:

Eric Chabrol. Caractérisation structurale et fonctionnelle d'une lectine de type-C des cellules de Langerhans : La Langérine. Sciences agricoles. Université de Grenoble, 2012. Français. NNT : 2012GREN022 . tel-00743636

**HAL Id: tel-00743636**

**<https://theses.hal.science/tel-00743636>**

Submitted on 19 Oct 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



## THÈSE

Pour obtenir le grade de

## DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Biologie Structurale et Nanobiologie**

Arrêté ministériel : du 7 août 2006

Présentée par

**Eric Chabrol**

Thèse dirigée par **Franck Fieschi**  
et codirigée par **Eric Girard**

préparée au sein de l'**Institut de Biologie Structurale**  
et de l'**Ecole Doctorale Chimie et Sciences du Vivant**

## Caractérisation structurale et fonctionnelle d'une lectine de type-C des cellules de Langerhans : La Langérine

Thèse soutenue publiquement le **29 mai 2012 à 14h**,  
devant le jury composé de :

**Mme Christelle Breton**

Professeur, Université de Grenoble, Présidente

**Mme Sylvie Ricard-Blum**

Professeur, Université de Lyon, Rapporteur

**M. Yves Bourne**

Professeur, Université de Marseille, Rapporteur

**Mme Dominique Durand**

Docteur, Université Paris-Sud, Examineur

**M. Franck Fieschi**

Professeur, Université de Grenoble I, Directeur de thèse

**M. Eric Girard**

Docteur, CEA Grenoble, Co-Directeur de thèse





## THÈSE

Pour obtenir le grade de

## DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Biologie Structurale et Nanobiologie**

Arrêté ministériel : du 7 août 2006

Présentée par

**Eric Chabrol**

Thèse dirigée par **Franck Fieschi**  
et codirigée par **Eric Girard**

préparée au sein de l'**Institut de Biologie Structurale**  
et de l'**Ecole Doctorale Chimie et Sciences du Vivant**

## Caractérisation structurale et fonctionnelle d'une lectine de type-C des cellules de Langerhans : La Langérine

Thèse soutenue publiquement le **29 mai 2012 à 14h**,  
devant le jury composé de :

**Mme Christelle Breton**

Professeur, Université de Grenoble, Présidente

**Mme Sylvie Ricard-Blum**

Professeur, Université de Lyon, Rapporteur

**M. Yves Bourne**

Professeur, Université de Marseille, Rapporteur

**Mme Dominique Durand**

Docteur, Université Paris-Sud, Examineur

**M. Franck Fieschi**

Professeur, Université de Grenoble I, Directeur de thèse

**M. Eric Girard**

Docteur, CEA Grenoble, Co-Directeur de thèse





# Remerciements

Je souhaite remercier le président de l'Université Joseph Fourier, le Collège Doctorale, l'École Doctorale de Chimie et Science du Vivant ainsi que l'Institut de Biologie Structurale pour l'attribution d'une bourse ministérielle et l'accueil au sein de l'institut.

Je dois une part importante à mon directeur de thèse Franck Fieschi pour sa confiance, son aide et son soutien au cours de cette thèse. Il m'a permis d'acquérir une grande autonomie au sein du laboratoire mais aussi de me permettre via son carnet d'adresse d'établir des collaborations qui ont permis d'augmenter la qualité de ce travail. Je souhaite aussi remercier mon co-directeur de thèse Éric Girard pour sa patience et son apprentissage de la cristallographie et la détermination de structure cristalline. Pour ce point notamment, je souhaite dédicacer ce travail à Richard Kahn, mon premier mentor pour la cristallographie qui est devenu un ami cher au fil des années.

La qualité de ce travail est dû en partie aux collaborations établies. Donc, un grand merci à Jenny, Colette, Dominique, Alessandra, Émile, Patrice, Antoine, Romain, Jesus et Pedro. Je tiens remercier aussi tout le réseau « carmusys » avec qui j'ai pu travailler.

Merci aussi aux membres de mon comité de thèse : Anne, Nicole, Jenny.

Merci au jury de thèse pour avoir passé le temps nécessaire à l'évaluation de ce travail et à leurs déplacements pour la soutenance.

Deux plateformes ont été utilisées au cours de ce travail, je souhaite donc les remercier. Merci donc à la plateforme Biacore et de cristallisation en nanogoutte.

Maintenant, une part important de mon temps s'est passée au laboratoire, j'ai donc créé des liens avec beaucoup de personnes. Je tiens à dédicacer ce travail à Corinne, Céline, Ieva, Isabelle, Sarah, Juliette, Lydia, Michel, Romain, Vincent, Gaël, Cristian, Nicolas. Merci aussi à toutes les personnes qui ont été d'un grand soutien, il y en a trop et j'aurais peur d'en oublier mais ils se reconnaîtront (LEM, Info, Secrétaire ...).

Enfin je souhaite dédier aussi ce travail à mon grand Pierre Chabrol décédé au cours de cette thèse. Tes valeurs, et ta présence resteront à jamais encrés en moi.



# Table des matières

<b>I</b>	<b>Introduction</b>	<b>25</b>
<b>1</b>	<b>Le système immunitaire</b>	<b>27</b>
1.1	Rôle général . . . . .	27
1.2	Système immunitaire inné . . . . .	28
1.3	Système immunitaire adaptatif . . . . .	30
1.4	Les Cellules Présentatrices d'Antigènes (APC) . . . . .	32
<b>2</b>	<b>Les cellules dendritiques (DC)</b>	<b>35</b>
2.1	Les différentes cellules dendritiques et leurs rôles . . . . .	37
2.2	La reconnaissance du « non-soi » par les cellules dendritiques : les « Toll-Like Receptors » et les lectines de type-C . . . . .	39
2.2.1	Les récepteurs de type Toll (TLR) . . . . .	39
2.2.2	Les lectines de type-C . . . . .	42
2.3	La maturation des cellules dendritiques . . . . .	42
2.4	Les cellules de Langerhans (LCs) . . . . .	43
2.4.1	La localisation des cellules de Langerhans . . . . .	44
2.4.2	Les différentes fonctions des cellules de Langerhans . . . . .	44
2.4.3	Le granule de Birbeck . . . . .	45
2.4.4	La reconnaissance de pathogène par les cellules de Langerhans . . . .	46
<b>3</b>	<b>Les cellules dendritiques et le VIH</b>	<b>49</b>
3.1	La trans-infection des lymphocytes T par les cellules dendritiques . . . . .	51
3.1.1	Processus de présentation du virus par le passage dans des phagosomes	52
3.1.2	La concentration du virus à la surface des DCs sans internalisation	52
3.1.3	La cis-infection . . . . .	53
3.2	Immuno-régulation par la reconnaissance du virus par DC-SIGN . . . . .	53
3.3	Accessibilité de DC-SIGN à la surface cellulaire comparée à CD4 et CCR5	53
3.4	Une défense naturelle contre le VIH : les cellules de Langerhans . . . . .	54

<b>4</b>	<b>Les lectines</b>	<b>59</b>
4.1	Informations générales sur les lectines . . . . .	59
4.1.1	Classification . . . . .	59
4.1.2	Reconnaissance des ligands . . . . .	60
4.2	La famille des lectines de type C . . . . .	62
4.2.1	Organisation structurale . . . . .	62
4.2.2	Spécificité pour la reconnaissance des saccharides . . . . .	63
4.2.3	Classification . . . . .	64
4.3	Les lectines de type C du groupe II . . . . .	65
4.3.1	Organisation structurale . . . . .	66
4.3.2	Le domaine coiled-coil . . . . .	66
<b>5</b>	<b>La Langérine</b>	<b>69</b>
5.1	Structure de Langérine . . . . .	69
5.1.1	Vue d'ensemble de la Langérine . . . . .	69
5.1.2	Implication de la Langérine dans le granule de Birbeck . . . . .	70
5.1.3	Structure du CRD . . . . .	71
5.1.4	Structure de l'ECD . . . . .	72
5.2	Reconnaissance et mode d'interaction de la Langérine . . . . .	73
5.2.1	Introduction et nomenclature usuelle des sucres . . . . .	73
5.2.2	Reconnaissance de la Langérine . . . . .	75
5.2.2.1	Reconnaissance des glucides de type mannose, N-acetyl- glucosamine et fucose . . . . .	76
5.2.2.2	Reconnaissance des saccharides sulfatés . . . . .	76
5.2.2.3	Reconnaissance des saccharides bactériens et de champi- gnons . . . . .	76
5.2.3	Mode d'interaction de la Langérine . . . . .	79
5.2.3.1	Interaction avec les sucres sulfatés. . . . .	79
5.2.3.2	Interaction avec le mannose . . . . .	79
5.2.3.3	Interaction avec le fucose . . . . .	81
<b>6</b>	<b>Les hydrates de carbone étudiés et leurs fonctions</b>	<b>85</b>
6.1	Les hydrates de carbone de type « High Mannose » . . . . .	85
6.2	Les glycosaminoglycanes (GAGs) . . . . .	86



<b>II</b>	<b>Les axes de travail</b>	<b>90</b>
<b>III</b>	<b>Matériel et méthodes</b>	<b>93</b>
<b>7</b>	<b>Biologie moléculaire</b>	<b>95</b>
7.1	Constructions plasmidiques utilisées . . . . .	95
7.2	Mutagenèse dirigée de Lg-S-CRD et Lg-ECD . . . . .	97
7.3	Transformation bactérienne . . . . .	97
<b>8</b>	<b>Biochimie, résonance plasmonique de surface et microscopie électro-</b>	<b>99</b>
	<b>nique</b>	
8.1	Expression et purification des protéines . . . . .	99
8.1.1	Expression et purification de Lg-S-CRD <sub>WT</sub> , Lg-S-CRD <sub>F241L</sub> et Lg-S-CRD <sub>K257A</sub> . . . . .	99
8.1.2	Expression et purification de Lg-ECD <sub>WT</sub> , Lg-ECD <sub>F241L</sub> , Lg-ECD <sub>K257A</sub> , Lg-ECD <sub>K299A</sub> , Lg-ECD <sub>K313A</sub> , Lg-ECD <sub>K299A K313A</sub> , Lg-S-ECD <sub>WT</sub> et Lg-S-CRD <sub>W264R</sub> . . . . .	100
8.2	La Résonance Plasmonique de Surface (SPR) . . . . .	101
8.2.1	Principe . . . . .	101
8.2.2	Fonctionnalisation des surfaces . . . . .	103
8.2.2.1	Surface d'étude de l'interaction entre la Langérine et la BSA glycosylée avec un trisaccharide composé de Man $\alpha$ 1-3Man $\alpha$ 1-6Man . . . . .	104
8.2.2.2	Surface d'étude de l'interaction entre la Langérine ECD et gp120 <sub>YU2</sub> (expériences complémentaires, chap. 12) . . . . .	104
8.2.2.3	Surface d'étude de l'interaction entre la Langérine et l'héparine (expériences complémentaires, cf chap. 12) . . . . .	105
8.2.2.4	Surface d'étude de l'interaction entre la Langérine ECD et la Langérine S-ECD (expériences complémentaires, chap. 12) . . . . .	105
8.2.3	Expériences d'inhibition . . . . .	106
8.2.3.1	Etude de l'inhibition de l'interaction entre la Langérine et l'héparine par des fragments d'héparine de différentes tailles et par l'héparine et l'héparane sulfate désulfaté en C6106 . . . . .	106
8.2.3.2	Etude de l'inhibition de l'interaction entre la Langérine et gp120 ou la BSA-Mann par l'héparine . . . . .	106
8.2.3.3	Etude de l'inhibition de l'interaction entre la Langérine et la BSA-Mann par les mono et trisaccharides sulfatés . . . . .	107

8.2.4	Microscopie électronique à coloration négative (expériences complémentaires, chap. 12)	107
<b>9</b>	<b>Cristallographie</b>	<b>109</b>
9.1	Le synchrotron	109
9.2	La cristallogenèse	110
9.3	Obtention des cristaux de Langérine CRD	111
9.4	Le trempage avec les sucres	112
9.5	Co-cristallisation	112
9.6	La cristallographie en biologie structurale	112
9.7	L'enregistrement des données	114
9.8	Protocole utilisé pour la détermination des structures	115
<b>IV</b>	<b>Résultats</b>	<b>119</b>
<b>10</b>	<b>Obtention des outils protéiques</b>	<b>121</b>
10.1	Expression et purification des Lg-S-CRDs	121
10.2	Expression et purification des Lg-ECDs	124
<b>11</b>	<b>Etude de l'effet des mutations F241L et W264R sur la Langérine et le granule de Birbeck</b>	<b>129</b>
<b>12</b>	<b>Etude de l'interaction de la Langérine avec des sucres de type glycosaminoglycanes</b>	<b>147</b>
12.1	Développement des expériences de SPR	148
12.1.1	Développement des surfaces	148
12.1.2	Amélioration de la régénération	148
12.1.3	Interaction non spécifique sur la piste témoin	149
12.2	Un nouveau mode d'interaction de la Langérine avec les glycosaminoglycanes	152
12.3	Résultats complémentaires	184
12.3.1	Détermination de la taille minimum requise de l'héparine par SPR	184
12.3.2	Le sulfate en C6 de la glucosamine important dans le mode d'interaction	187
12.3.3	Expérience d'inhibition de l'interaction Langérine/gp120 ou BSA-Man par l'héparine	188
12.3.4	Résultats préliminaires sur la formation de complexe macromoléculaire formé de Langérine et d'héparine : vers une description des granules de Birbeck.	189

12.3.4.1	Microscopie électronique . . . . .	189
12.3.4.2	Propriétés d'assemblage de la Langérine . . . . .	191
<b>13</b>	<b>Caractérisation structurale de l'interaction de la Langérine avec des sucres sulfatés</b>	<b>193</b>
13.1	Introduction . . . . .	193
13.2	Etude fonctionnelle par Résonance plasmonique de surface . . . . .	197
13.3	Etude structurale de la Langérine avec les sucres de type GlcNac, GlcNac6S et GlcNS6S . . . . .	201
13.4	Etude par RMN de l'interaction avec des trisaccharides sulfatés . . . . .	206
13.5	Conclusion . . . . .	211
<b>14</b>	<b>Développement d'inhibiteurs de DC-SIGN</b>	<b>215</b>
<b>V</b>	<b>Discussion</b>	<b>237</b>
<b>VI</b>	<b>Bibliographie</b>	<b>249</b>
<b>VII</b>	<b>Annexes</b>	<b>265</b>

---

# Table des figures

1.1	Schéma représentant les principaux constituants cellulaires et solubles de l'immunité innée (cercle bleu) et de l'immunité adaptative (cercle rouge).[30]	29
1.2	Les voies du système du complément. (cellular and molecular immunology, fifth edition) . . . . .	30
1.3	Vue d'ensemble des branches humorale et à médiation cellulaire du système immunitaire. (adapté à partir de « Immunologie » de Janis Kuby). . . . .	32
1.4	Représentation du pH phagosomal en fonction du temps pour les neutrophiles (A), les macrophages (B) et pour les cellules dendritiques (C). V-ATPase est une pompe à proton permettant l'acidification du phagosome. NOX2 est une isoforme de la NADPH oxydase.[87] . . . . .	34
2.1	Micrographie de peau indiquant les deux couches principales de cellules. [83]	35
2.2	Localisation et migration de DCs dans le cas d'une infection par le VIH.[115]	36
2.3	Image de cellules dendritiques. <a href="http://www.uni-ulm.de/expane/images/dendritic.jpg">http://www.uni-ulm.de/expane/images/dendritic.jpg</a> . . . . .	37
2.4	Représentation schématique des sous-populations de cellules dendritiques et de leurs progéniteurs.[4] . . . . .	38
2.5	Représentation des différents TLRs, leurs ligands et les voies de signalisation impliquées. dsRNA : ARN double brin. ssRNA : ARN simple brin. IRF3 : facteur de transcription de l'interféron. NFκB : facteur de transcription. Image réalisée par l'entreprise ABD SEROTEC <a href="http://www.abdserotec.com/minireview_toll_like_receptors_tlrs_-1060.html">http://www.abdserotec.com/minireview_toll_like_receptors_tlrs_-1060.html</a> .	40
2.6	Maturation des cellules dendritiques. A gauche sont montrés les différents facteurs déclenchant la maturation. A droite sont représentées les principales propriétés de chaque étape de la maturation/différentiation.[4]. traduit par <a href="http://www.sante.univ-nantes.fr">www.sante.univ-nantes.fr</a> . . . . .	43

2.7	Micrographie d'une cellule de Langerhans (A) et d'un agrandissement sur un granule de Birbeck (B). G : granule de Birbeck, K : keratinocyte adjacent, L : lysosome, n : noyau, X : organite rond avec une structure interne granulaire.[8] . . . . .	44
2.8	Modèles de coupe de granule de Birbeck dans différents plans.[84] . . . . .	45
2.9	L'expression de Langérine recombinante induit la formation de BG dans des fibroblastes murins.[104] . . . . .	46
2.10	Reconnaissance de champignons par la Langérine. [23] . . . . .	47
3.1	Schéma d'une particule de VIH. [36] . . . . .	49
3.2	Schéma d'un mode d'infection du VIH. [46] . . . . .	50
3.3	Les différents processus observés après l'attachement du VIH par DC-SIGN. [102] . . . . .	51
3.4	Interaction DCs dermiques avec le VIH. Interaction du VIH avec les DCs dermiques via DC-SIGN : i) internalisation et dégradation du virus dans les lysosomes. ii) internalisation du virus dans des corps multivésiculaires conduisant à la présentation du virus intact à la surface cellulaire. iii) dégradation du virus par le protéasome. iv) infection directe des DCs par le virus (cis-infection).[26] . . . . .	52
3.5	A : Schéma représentant les tailles respectives des différents récepteurs, DC-SIGN, CD4 et CCR5. La structure utilisée pour DC-SIGN résulte d'une étude par SAXS [97], celle de CCR5 est un modèle (1ND8). B : Combinaison d'une topographie (gris) et d'images de fluorescences (couleur) [19]. La fluorescence représente les molécules de DC-SIGN à la surface cellulaire d'une cellule dendritique. . . . .	54
3.6	Clichés de microscopie électronique de coupe de cellules de Langerhans en présence de VIH. La Langerine et la protéine p24 de capsid du VIH sont marquées par des anticorps couplés avec des billes d'or de 10 nm et 15 nm respectivement. A) Co-localisation de la Langerine et du virus à la surface des cellules de Langerhans. B) Internalisation du virus dans les granules de Birbeck. La barre d'échelle représente 100 nm. [24] . . . . .	55
3.7	Représentation schématique de l'interaction du virus avec de LCs en condition "normale" (A) et lorsque la Langérine est inhibée. [24] . . . . .	56

3.8	Interaction DCs dermiques avec le VIH. a) Interaction du VIH avec les cellules de Langerhans : i) internalisation du virus dans les granules de birbeck qui permet de protéger l'infection. ii) L'inhibition de la Langérine (ou la saturation de la Langérine) conduit à l'infection des cellules de Langerhans qui permettra par le virion nouvellement formé d'infecter les lymphocytes T. b) Interaction du VIH avec les DCs dermiques via DC-SIGN : i) internalisation et dégradation du virus dans les lysosomes. ii) internalisation du virus dans des corps multivésiculaires conduisant à la présentation du virus intact à la surface cellulaire (Trans-infection). iii) dégradation du virus par le protéasome. iv) infection directe des DCs par le virus (Cis-infection).[26]	57
4.1	Représentation schématique des plusieurs modes d'avidités engendrés par les lectines. A : multivalence de surface. B : multivalence par la présence de plusieurs CRDs au sein d'une même chaîne. C : multivalence par oligomérisation de la protéine. D : association de la multivalence de surface et de l'oligomérisation. . . . .	60
4.2	A : représentation de la structure secondaire des CTLDs. Les ponts disulfures possibles dans ce type de lectine sont représentés par les lignes numérotées. B : Structure cristallographique de DC-SIGN. En bleus sont représentés les ponts disulfures (PDB id : 2XR6). En verts les ions calciums. <a href="http://www.imperial.ac.uk/research/animallectins/ctld/disulfides.html">http://www.imperial.ac.uk/research/animallectins/ctld/disulfides.html</a> . . . . .	62
4.3	Alignement des séquences protéiques des CRDs de la Langérine, DC-SIGN, DC-SIGNR et la MBP. En vert, résidues des brins $\beta$ . En rouge, résidues des hélices $\alpha$ . Les lignes jaunes représentent les ponts disulfures. Le surlignage gris montre les résidues conservés chez ces lectines en interaction directe avec l'atome de calcium du site 1. Le carré bleu met en évidence le site 1. [100] . . . . .	63
4.4	Structure cristallographique de lectine de type-C possédant le motif EPN (A) et QPD (B). A : structure cristallographique de la Langérine en présence de maltose (3P7H) [14]. B : structure cristallographique du récepteur scavenger en présence de galactose (2OX9) [37]. . . . .	64
4.5	Les différents groupes de lectines de type C. Image réalisée par le Prof. Driickamer <a href="http://www.imperial.ac.uk/research/animallectins/ctld/lectins.html">http://www.imperial.ac.uk/research/animallectins/ctld/lectins.html</a> . . . . .	65

4.6	Représentation schématique de l'organisation générale des lectines de type C du groupe II. CD : domaine cytoplasmique (cytoplasmic domain) ; TM : domaine transmembranaire (transmembrane domain) ; ND : domaine neck (neck domain) ; CRD : domaine de reconnaissance des sucres ; ECD : domaine extracellulaire composé du domaine neck et du CRD. . . . .	66
4.7	Représentation d'un heptade d'acides aminés formant une hélice $\alpha$ de type coiled-coil. A : vue du dessus. B : vue latérale. C : schéma de distribution des acides aminés d'une heptade. . . . .	67
4.8	Quatre modèles d'organisation oligomérique de type coiled-coil parallèle avec des exemples de structures résolues. A : Un dimère de GCN4 (PDB : 2ZTA, [78]). B : Un trimère de MBP-A (PDB : 1RTM, [110]). C : Un tétramère de Tetrabrachion (PDB : 1FE6, [96]). . . . .	68
5.1	Schéma représentant l'architecture de la Langérine. CD : domaine cytoplasmique. TMD : domaine transmembranaire. ND : domaine neck. CTLD : domaine de lectine de type-C. ECD : domaine extracellulaire. [106] . . . .	69
5.2	Micrographies des différentes constructions de la Langérine. A : transfection réalisée avec le gène de la Langérine déplété du CRD, B : transfection réalisée avec le gène de la Langérine déplété du domaine cytoplasmique, C : transfection réalisée avec le gène de la Langérine contenant des mutations dans le domaine cytoplasmique. [100] . . . . .	70
5.3	Les granules de Birbeck dans les cellules de Langerhans et les modèles de l'organisation macromoléculaire. A : Invagination de la membrane plasmique formant une structure en sandwich de membrane dans une cellule de Langerhans. B : Granules de Birbeck des cellules de Langerhans. C : Deux modèles de l'organisation macromoléculaire de la Langérine dans le granule de Biebeck. D : Modèle de la déformation de la membrane plasmique par la Langérine. [100] . . . . .	71
5.4	Structure cristallographique du CRD de la Langérine (A) alignée avec celles de DC-SIGN (vert), de la MBP (orange) et de DC-SIGNR (cyan) (B). C : Zoom sur l'environnement du site calcium 2. D : Zoom du site calcium 1. E : Zoom du site 4. . . . .	72
5.5	Structure cristallographique d'une partie du domaine extracellulaire de la Langérine (PDB : 3KQG). A : coiled-coil vue du dessus et de côté. B : Interface ND d'un monomère avec le CRD du monomère voisin. C : interface CRD/CRD. D : interface ND et CRD d'un même monomère. [38] . . . .	73
5.6	Cyclisation du D-glucose. . . . .	74
5.7	Nomenclature proposée par le consortium de glycobiologie. . . . .	75



5.8	Interaction de Langérine fluorescente avec une puce à glycanes. Les barres vertes représentent les structures oligosaccharidiques ou polysaccharidiques de type « High-Mannose ». Les barres bleues sont les oligosaccharides terminant par un N-acetyl-glucosamine. Les rouges sont les oligosaccharides terminant par un fucose. Les barres violettes représentent les oligosaccharides terminant par un galactose sulfaté en position 6 [38]. . . . .	77
5.9	Interaction de la Langérine avec une puce fonctionnalisée de polysaccharides bactériens [39]. . . . .	78
5.10	Spécificité de reconnaissance de DC-SIGN (haut) et la Langérine (bas). [53]	78
5.11	Structure cristallographique du CRD de la Langérine en présence d'un 6 sulfogalactose $\beta$ 1-4 N-acetylglucosamine. a : Vue de la surface de la Langérine en interaction avec le disaccharide. b : Zoom sur le mode d'interaction du disaccharide et de sa densité électronique avec la Langérine. c : Zoom de l'interaction du galactose sulfaté avec l'ion calcium. d : comparaison de l'orientation du galactose sulfaté et du mannose en interaction avec l'ion calcium. e : Implication de l'hydroxyle en C2 et de la liaison osidique dans l'interaction. f : Zoom sur le réseau d'interaction engendré par le sulfate en position 6 du galactose. [39] . . . . .	80
5.12	Interaction du CRD de la Langérine avec un disaccharide Man $\alpha$ 1-2Man. A : Orientation mettant en évidence les deux orientations possibles de mannose en interaction tout en gardant une superposition parfaite des hydroxyles en C3 et C4. B : Implication de la lysine 299 dans les deux orientations, soit en interaction avec l'hydroxyle anomérique soit avec celui en C2. C : Mode d'interaction avec le Man $\alpha$ 1-2Man lorsque le mannose de l'extrémité non réductrice est en interaction avec le calcium. D : Mode d'interaction avec le Man $\alpha$ 1-2Man lorsque le mannose de l'extrémité réductrice est en interaction avec le calcium.[39] . . . . .	81
5.13	Interaction du CTLD de la Langérine avec le trisaccharide du groupe sanguin B. a-f : différentes orientations de l'interaction de la Langérine avec le trisaccharide du groupe sanguin B. g : superposition du lewis x et du trisaccharide avec le mode d'interaction de la Langérine. h : superposition du lewis x et du trisaccharide avec le mode d'interaction de DC-SIGN. i : orientation du fucose en interaction via les hydroxyles en C2 et C3 (jaune) ou avec les hydroxyles en C3 et C4 (mauve). [39] . . . . .	83
6.1	Schéma représentatif des structures de type « High-Mannose ». . . . .	86
6.2	Représentation schématique des différents glycosaminoglycanes retrouvés. [45] . . . . .	87

6.3	Illustration de la composition des différents groupes de glycosaminoglycanes. Les différentes modifications possibles sont représentées en rouge, vert et jaune pour les différentes sulfotransférases et en bleu pour les épimérases. Le dermatane sulfate correspond à la chondroïtine sulfate B. [9]	88
6.4	Représentation des différentes modifications pouvant être effectuées sur les chaînes d'héparane sulfate révélant la diversité des structures saccharidiques. Les modifications sont réalisées par des épimérases et des sulfotransférases dans l'appareil de Golgi. [9]	89
6.5	Modification de la structure du granule de Birbeck par la mutation F244L et perte de la formation du granule par la mutation W264R. A : fibroblaste transfecté avec le gène de la Langérine de souris. B : Fibroblaste transfecté avec le gène de la Langérine de souris contenant la mutation F244L[105]. C : Cellule de Langerhans humaine contenant le gène de la Langérine avec la mutation W264R [107].	92
7.1	Carte du plasmide pET30b.	96
8.1	Constructions protéiques obtenues.	99
8.2	Schéma de la technique de résonance plasmonique de surface, de la variation de l'angle de résonance et d'un sensorgramme montrant les différentes phases d'une interaction. [17]	102
8.3	Couplage covalent par les amines protonées d'une biomolécules. (note d'application GE Healthcare)	103
9.1	Schéma représentant les équipements de bases du synchrotron SOLEIL (Paris). Source : Synchrotron SOLEIL	110
9.2	Schéma représentant le diagramme de phase (haut) pour la méthode de goutte suspendue (bas). D'après le livre « Cristallographie » réalisé par Dieter Schwarzenbach et Gervais Chapuis.	111
9.3	Exemple de motif de diffraction pour le lysosyme. Cliché réalisé par Romain Talon (IBS).	113
10.1	SDS-PAGE 15% montrant l'expression de Lg-S-CRD. M : marqueur de poids moléculaire. I <sub>0</sub> : Extrait de bactéries avant induction. I <sub>∞</sub> : Extrait de bactéries à la fin de l'expression de Lg-S-CRD.	122
10.2	Chromatogramme de purification de Lg-S-CRD. Bleu : Absorbance à 280 nm. Vert : Concentration en % du tampon d'élution (axe des ordonnées secondaires).	123

10.3	A : Chromatogramme de purification sur colonne de mannose. Noir : échantillon concentré de Lg-S-CRD. Gris : contrôle négatif (BSA). B : SDS-PAGE 15% des protéines purifiées. M : Marqueur de poids moléculaires. 1 : extrait purifié de Lg-S-CRD. . . . .	123
10.4	SDS-PAGE 12% de l'expression de Lg-ECD. M : Marqueur de poids moléculaires. 1 : Extrait total avant induction. 2 : Extrait total à la fin de l'induction . . . . .	124
10.5	Gel SDS-PAGE 12% des étapes de renaturation des protéines. M : Marqueur de poids moléculaires. 1 : Extrait total. 2 : Surnageant après le premier lavage des corps d'inclusion (tampon : Tris-HCl pH8 25 mM; NaCl 150 mM). 3 : Surnageant après le deuxième lavage des corps d'inclusion (tampon : Tris-HCl pH8 25 mM; NaCl 150 mM; urée 2 M; triton X100 1%). 4 : Surnageant après le troisième lavage des corps d'inclusion (tampon : Tris-HCl pH8 25 mM; NaCl 150 mM). 5 : 'Fraction soluble des corps d'inclusion solubilisés (tampon : Tris-HCl pH8 25 mM; NaCl 150 mM, chlorure de guanidinium 6 M; $\beta$ -mercaptoethanol 0,01%). . . . .	125
10.6	Chromatogramme de purification de Lg-ECD sur une colonne de Mannan-Agarose. Vert : pourcentage de tampon B. Rouge : conductivité. . . . .	126
10.7	Donnée de sédimentation obtenue par ultracentrifugation analytique de Lg-ECD [100]. Deux concentrations de Lg-ECD ont été analysées, 1 mg/ml (pic du premier plan) et 0,1 mg/mL (pic du second plan). $S_{20,w}$ : coefficient de sédimentation. $C(s)$ : concentration normalisée de l'espèce en solution. . . .	127
12.1	Représentation de la réponse maximum, normalisée par rapport au RU, en fonction du nombre de cycles pour une concentration fixe de Lg-ECD (500 nM). Rouge : effet de la régénération par le SDS. Bleu : effet de la régénération par le $MgCl_2$ . . . . .	149
12.2	Sensorgrammes représentant l'interaction de Lg-ECD à 2 $\mu$ M sur la surface témoin (streptavidine) et la surface d'étude (streptavidine + héparine) en tampon de course calcium (A) ou EDTA (B). C : Histogrammes des pourcentages de l'interaction non désirée par rapport au signal total pour différentes concentrations en Langérine ECD en tampon calcium (gris) et en tampon EDTA (bleu). . . . .	150
12.3	Sensorgrammes de la greffe de la streptavidine précédée du pré-traitement (A) et de l'interaction non-désirée de la Langérine à une concentration de 500 nM avec ou sans le pré-traitement (B). . . . .	151

12.4	Schéma de la digestion enzymatique de l'héparine par l'héparinase III. Modification de l'acide glucuronique par la présence d'une double liaison entre les carbones 4 et 5. [117] . . . . .	184
12.5	Sensorgrammes de l'expérience d'inhibition de l'interaction de la Langérine avec l'héparine par différentes tailles de fragments d'héparine. Débit de 10 $\mu$ L/min. Injection de 50 $\mu$ L d'échantillons composés de Langérine à 100 nM et des différents fragments à 1 $\mu$ M. Les échantillons sont préparés dans le tampon de course (25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $\text{CaCl}_2$ et de 0,005% de surfactant P20). . . . .	185
12.6	Sensorgrammes de l'expérience d'inhibition avec les héparines et héparanes sulfates traités par HSulf2 ou non. . . . .	186
12.7	Représentation de la chaîne d'héparine. Le cadre rouge met en évidence le disaccharide ciblé dans la désulfatation par l'enzyme HSulf2. Le cadre bleu montre le sulfate enlevé par HSulf2. . . . .	187
12.8	Sensorgrammes de l'expérience d'inhibition de l'interaction Langérine/gp120 (A) ou BSA-Man (B) par l'héparine. . . . .	188
12.9	Micrographies de la Langérine ECD seule (A,B) ou en complexe avec de l'héparine (C, D, E, F, G). A-B : LgECD seule. C-D-E-F-G : LgECD en complexe avec l'héparine 15 kDa pour une ratio 1 :1. L'échelle est identique pour toutes les micrographies. . . . .	190
12.10	Sensorgramme de l'expérience d'injection successive de la Langérine et de l'héparine. A : Injection de la Langérine seule (50 $\mu$ L à 500 nM). B : Injection de l'héparine (100 $\mu$ L à 100 nM) puis de la Langérine (50 $\mu$ L à 500 nM). C : Injection de la Langérine puis de l'héparine. Le débit est de 20 $\mu$ L/min dans le tampon calcium (25 mM Tris pH 8, 150 mM NaCl, 4 mM $\text{CaCl}_2$ et 0,005% P20). . . . .	191
13.1	Structure du tris-dipicolinate de lanthanide (A) et de la capacité de pontage de cette molécule observée dans le cas du lysosyme [81]. . . . .	194
13.2	Structures des saccharides utilisés. . . . .	195
13.3	Sensorgramme des immobilisations de BSA (A) et BSA-Man (B). . . . .	198
13.4	Sensorgrammes d'une expérience d'inhibition avec le GlcNac6S (A) et la représentation graphique permettant d'estimer l' $\text{IC}_{50}$ (B). . . . .	199
13.5	Histogramme des $\text{IC}_{50}$ des sucres de type GlcNac, GlcNac6S, GlcNS6S et Trila pour inhiber l'interaction de la Langérine ou DC-SIGN avec la BSA-Man. . . . .	200
13.6	Structures cristallographiques de la Langérine (A, code PDB : 3C22) et de DC-SIGN (B, code PDB : 2IT5). . . . .	201

---

13.7	Carte de densité électronique $2F_O-F_C$ au niveau des saccharides (contour : 1 sigma). A : GlcNac. B : GlcNac6S. C : GlcNS6S. . . . .	203
13.8	Structure cristallographique mettant en évidence l'orientation des saccharides au niveau du site calcium. A-B : GlcNac. C-D : GlcNac6S. E-F : GlcNS6S. Les pointillés représentent le réseau de liaisons impliquées dans l'interaction (distance entre 2,5 et 3,5 Å). . . . .	205
13.9	Schéma basique d'une expérience de STD. . . . .	206
13.10	Représentation des deux séries de trisaccharides de synthèse produites par nos collaborateurs. . . . .	207
13.11	Structure du trisaccharide Tri1a et Tri1b. . . . .	208
13.12	Spectre 1D du proton de l'expérience STD avec le Tri1a en calcium (A) et en EDTA (B) et avec le Tri1b en calcium (C) et en EDTA (D). Les points rouges mettent en évidence les protons pour lesquels il y a eu échange de saturation. C : le spectre du haut est le spectre de référence et celui du bas est le spectre après soustraction avec le spectre STD. D : le spectre du haut est le spectre de référence et celui du bas est le spectre STD. . . . .	209
13.13	Quantification du transfert de saturation des protons sur le Tri1a (A) et le Tri2b (B). . . . .	210
13.14	Structure de l'hexasaccharide utilisé. . . . .	211
13.15	Structure cristallographique du CRD de la Langérine en présence de mannose (ID : 3P5F). . . . .	211
14.1	Combinaison du modèle de l'organisation de la Langérine en alternance dans les BGs et de la modélisation avec l'héparine. . . . .	245
14.2	Combinaison du modèle de l'organisation de la Langérine tête bêche dans les BGs et de la modélisation avec l'héparine. . . . .	245

# Abréviations

ADN :	Acide désoxyribonucléique
APC :	Antigen-Presenting Cell (cellule présentatrice d'antigène)
ARN :	Acide ribonucléique
BCR :	B cell receptor (récepteur des cellules B)
BG :	Birbeck Granule (granule de Birbeck)
BSA :	Bovin serum albumin (albumine de sérum bovin)
BSA-Man :	BSA glycosylée en 12 sites avec un trisaccharide composé de $\text{Man}\alpha 1\text{-3Man}\alpha 1\text{-6Man}$
$\text{CaCl}_2$ :	Chlorure de calcium
CD :	Cytoplasmic Domain (domaine cytoplasmique)
CFG :	Consortium for Functional Glycomic
CMH :	Complexe Majeur d'Histocompatibilité
CMS :	Structure de type sandwich de cytomembrane
CRD :	Carbohydrate Recognition Domain (domaine de reconnaissance des hydrates de carbone)
CS :	Chondroïtine Sulfate
CTLD :	C-Type Lectin Domain
Da :	Dalton
DC :	Dendritic Cell (cellule dendritique)
DC-SIGN :	Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin
dNTP :	Mélange des quatres désoxyribonucléotides
$\text{EC}_{50}$ :	Concentration de protéine pour laquelle la réponse représente 50% du signal maximum
ECD :	Extracellular Domain (domaine extracellulaire)
EDC :	1-ethyl-3-(3-diméthylaminopropyl) carbodiimide

---

EDTA :	Acide éthylène diamine tétraacétique
ESRF :	European Synchrotron Radiation Facility
F241L :	Mutation de la phénylalanine à la position 241 par une Leucine de la Langérine humaine
F244L :	Mutation de la phénylalanine à la position 241 par une Leucine de la Langérine murine
GAG :	Glycosaminoglycane
Gal :	Galactose
GlcNac :	N-acétyl-glucosamine
GlcNac6S :	6-sulfo, N-acétyl-glucosamine
GlcNS6S :	N, 6 disulfo-glucosamine
gp120 :	Glycoprotéine d'enveloppe 120 du virus de l'immunodéficience humaine
HBS-P :	Hepes 100 mM pH 7,4 ; NaCl 150 mM ; Surfactant P20 0,005%
HCl :	Acide chloridrique
Hepes :	Acide 4-(2-hydroxyéthyl)-1-pipérazine éthane sulfonique
Hp :	Héparine
Hs :	Héparane sulfate
IBS :	Institut de Biologie Structurale
IC <sub>50</sub> :	Concentration en inhibiteur pour induire 50% d'inhibition
Idoa :	Acide iduronique
IFN :	Interféron
IL :	Interleukine
IPC :	Interferon Producing Cell (cellule productrice d'interféron)
IPTG :	Isopropyl $\beta$ -D-1-thiogalactopyranoside
K <sub>d</sub> :	Constante de dissociation
KS :	Kératane sulfate
LB :	Milieu Luria-Bertani
LC :	Langerhans Cell (cellule de Langerhans)
LPS :	Lipopolysaccharide
Man :	Mannose
MBP :	Maltose Binding Protein

---

mDC :	Myeloid Dendritic cell (cellule dendritique myéloïde)
MES :	Acide 2-morpholino éthanesulfonique monohydraté
MgCl <sub>2</sub> :	Chlorure de magnésium
MGL :	Macrophage Galactose Lectin
moDC :	Monocyte derived Dendritic cell (cellule dendritique dérivée de monocyte)
NaCl :	Chlorure de sodium
ND :	Neck Domain (domaine neck)
NF $\kappa$ B :	Nuclear factor kappa-light-chain-enhancer of activated B cells
NHS :	N-Hydroxysuccinimide
NK :	Natural Killer
NOX <sub>2</sub> :	NADPH oxydase
PAMP :	Pathogen Associated Molecular Pattern (motif moléculaire associé au pathogène)
PCR :	Polymerase Chain Reaction
PCS :	Site de clonages multiples
pDC :	Plasmacytoid Dendritic cell (cellule dendritique plasmacytoïde)
PEG :	Polyéthylène glycol
PRR :	Pattern Recognition Receptor
RMN :	Résonance Magnétique Nucléaire
RU :	Resonance Unit (unité de résonance)
SAXS :	Small Angle X-ray Scattering
SDS :	Dodécylsulfate de sodium
SDS-PAGE :	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SPR :	Surface Plasmon Resonance (résonance plasmonique de surface)
STD :	Saturation Transfer Difference (différence de transfert de saturation)
TCR :	T Cell Receptor (récepteur des cellules T)
TLR :	Toll Like Receptor (récepteur de type Toll)
TMD :	Transmembrane Domain (domaine transmembranaire)
TNF :	Tumor Necrosis Factor (facteur de nécrose tumorale)
Tris :	2-amino-2-hydroxyméthyl-1,3-propanediol
VIH :	Virus de l'Immudéficience Humaine
W264R :	Mutation du tryptophane à la position 264 par une argine de la Langérine humaine



# Préambule

---

Ce travail de thèse s'est focalisé sur la caractérisation structurale et fonctionnelle de la Langérine. Cette protéine appartient à la famille des lectines de type-C. Cette famille de protéine est caractérisée par la présence d'un domaine de reconnaissance des sucres (CRD, « Carbohydrate Recognition Domain »). Cette capacité de reconnaissance permet de reconnaître un nombre important de pathogènes comme des bactéries, champignons ou encore virus. De plus, cette protéine est principalement exprimée dans un sous-type de cellules dendritiques, les cellules de Langerhans. Ces cellules sont fortement impliquées dans l'immunité innée comme acquise. L'implication de ces cellules au niveau du système immunitaire est dépendant de l'expression des récepteurs à la surface cellulaire. Ainsi, les cellules de Langerhans permettent notamment de protéger l'organisme face à une infection par le VIH *via* la reconnaissance du virus par la Langérine et son internalisation dans un organite intracellulaire, le granule de Birbeck. Le granule de Birbeck est un organite spécifique des cellules de Langerhans et sa formation est directement sous la dépendance de l'expression de la Langérine.

Au début de cette thèse, peu de choses étaient connues des capacités de reconnaissances de la Langérine, de son organisation au sein du granule de Birbeck ainsi que des mécanismes de formation de cet organite. Dans le contexte d'un développement croissant des recherches dans le domaine des interactions hôtes-pathogènes, la caractérisation des spécificités de reconnaissances de la Langérine ainsi que des mécanismes associés de capture et d'internalisation des pathogènes revêtaient un intérêt croissant. En effet, une meilleure connaissance des fonctions de la Langérine et de son rôle dans les cellules de Langerhans pourrait ouvrir des perspectives dans des approches de vaccination ou de régulation du système immunitaire *via* ces cellules.

Nous avons donc divisé l'étude en deux axes majeurs :

1. L'étude de l'implication de la Langérine et de sa structure dans la formation et la structure des granules de Birbeck. Pour cela, nous avons voulu caractériser l'effet moléculaire de deux mutations décrites pour provoquer la perte des granules de Birbeck ou la modification de ceux-ci. Pour cela, nous avons combiné des approches de biochimie, de biologie cellulaire et plus particulièrement de biophysiques telles que le dichroïsme circulaire, la cristallographie ou encore la diffusion des rayons X aux petits angles et la résonance plasmonique de surface.
2. Le deuxième axe de travail est focalisé à la caractérisation des capacités d'interactions de la Langérine et l'étude d'une nouvelle. Cet axe est divisé en deux sous parties :
  - L'étude de l'interaction entre la Langérine et les glycosaminoglycanes. Cette orientation vient de l'analyse de données bibliographique publiées en début de thèse par un groupe japonais. Cete équipe a mis en évidence la capacité d'interaction avec des

---

sucres sulfatés et a suggéré la possibilité d’une interaction entre la Langérine et le kératane sulfate. Ce polymère saccharidique appartient à la famille des glycosaminoglycanes et nous nous sommes donc focalisés sur la caractérisation de l’interaction entre la Langérine et ce type particulier de sucres qui sont des constituants des matrices extracellulaires. Cette étude a été réalisée principalement en utilisant la résonance plasmonique de surface et la modélisation.

- Au cours de l’étude précédente nous avons observé un mode d’interaction différent entre les glycosaminoglycanes et leurs monosaccharides associés. Par conséquent, cette partie réalisée en fin de thèse se consacre à la caractérisation fonctionnelle, par la résonance plasmonique de surface, et structurale, par la cristallographie, de l’interaction entre la Langérine et des monosaccharides de type glucosamine sulfaté. Les informations obtenues sur les modes d’interactions seront discutées dans le contexte du développement d’inhibiteurs spécifiques des récepteurs lectines de cette famille.

Ainsi du fait du thème de travail, ce manuscrit de thèse commencera par une introduction rappelant les bases du fonctionnement du système immunitaire. Cela sera ensuite suivi d’un chapitre détaillant un peu plus les cellules dendritiques et leurs rôles dans l’activation d’une réponse immunitaire spécifique. Le cas du VIH sera plus spécifiquement abordé et cette introduction continuera sur une présentation générale de la famille des récepteurs lectines de type-C étudiés et plus particulièrement l’état des connaissances sur la Langérine, membre de cette famille. Du fait, d’une des études menées dans la thèse, un chapitre rappelant les structures et les propriétés des glycosaminoglycanes (oligosaccharides sulfatés de la matrice extracellulaire) clôturera cette introduction bibliographique.

---

# Première partie

## Introduction

---

# Chapitre 1

## Le système immunitaire

Le système immunitaire est l'ensemble d'un réseau étroitement lié d'éléments (organes, cellules, molécules inflammatoires...) participant à la reconnaissance du « soi » et à sa tolérance ainsi qu'à la défense contre le « non-soi » (bactéries, virus, parasites, cellules du soi altérées...).

### 1.1 Rôle général

Dès la naissance, nous sommes constamment en présence de micro-organismes capables de provoquer des effets indésirables dans notre organisme. Sans mécanismes de protection, nous succomberions rapidement à des maladies infectieuses provoquées par toutes sortes de microbes. Pour cela, les mammifères ont développé un réseau extrêmement complexe de défenses permettant le maintien de l'intégrité physique de l'organisme. Ces divers mécanismes de défense sont regroupés sous deux grands types d'immunités :

Le premier, répondant à une nécessité de combattre rapidement tous types d'éléments infectieux, est non spécifique mais efficace contre un grand nombre de pathogènes. Ce sont les mécanismes de l'immunité innée.

Lorsque ceux-ci sont dépassés et ne contiennent pas la menace, le second type « d'immunité », appelé immunité adaptative (ou acquise), met en place des réponses spécifiques au pathogène considéré après un temps d'adaptation.

Ces deux types d'immunité sont complémentaires et permettent une réponse efficace contre tout élément étranger (bactéries, virus, transplants, toxines...).

Dans les parties suivantes, l'immunité innée et adaptative seront développées séparément. Les différents types cellulaires et leurs rôles dans le maintien de l'intégrité de l'organisme seront évoqués. Par la suite, nous nous focaliserons sur des cellules particulières du système immunitaire. Ces cellules permettent de constituer un lien fort entre les deux types d'immunité, elles sont appelées les cellules présentatrices d'antigènes (APC,

« Antigen Presenting Cell »).

## 1.2 Système immunitaire inné

Comme décrit précédemment, l'immunité innée permet de répondre très rapidement à un grand nombre de pathogènes avec une spécificité large et peu restrictive. Pour remplir ce rôle, l'organisme a développé différentes barrières défensives : des barrières physiques, comme la peau et les muqueuses qui protègent efficacement contre l'entrée de pathogènes. Des barrières physiologiques, comme l'acidité de l'estomac qui empêche la plupart des micro-organismes de vivre dans ce milieu. Une barrière « inflammatoire » qui permet, par exemple, d'amener un grand nombre de cellules immunitaires à la zone d'infection. Et enfin, une barrière phagocytaire permettant l'internalisation de micro-organismes dans des compartiments cellulaires pour leurs dégradations.

Par la suite, nous allons plus particulièrement nous intéresser à la composante cellulaire et phagocytaire de cette immunité. Mais avant de rentrer dans les détails, la description des cellules impliquées dans cette immunité est nécessaire. On retrouve différents types de cellules, les granulocytes (neutrophiles, basophiles et les éosinophiles), les macrophages, les mastocytes et les cellules dendritiques (Figure 1, cercle bleu).

Cette multitude de différents types cellulaires correspond à différentes fonctions, pour répondre efficacement à une infection, mais aussi à différentes localisations, tissulaires ou circulantes.

Dans le sang, on retrouve principalement les granulocytes. Parmi ceux-ci, on trouve trois sous-types cellulaires ayant des fonctions différentes :

- Les neutrophiles qui sont des phagocytes, c'est-à-dire qu'elles sont capables de reconnaître, internaliser et dégrader les pathogènes reconnus.
- Les basophiles qui sécrètent de l'histamine permettant d'activer la réaction inflammatoire.
- Les éosinophiles qui permettent de réguler par sécrétion de molécules la réaction inflammatoire activée par les basophiles, mais aussi de détruire les pathogènes en sécrétant des enzymes.

Dans les tissus, on retrouve des cellules phagocytaires : les macrophages et les cellules dendritiques. Ces cellules sont capables de reconnaître les molécules du « non-soi », de les internaliser, les dégrader et de présenter des antigènes aux cellules du système immunitaire adaptatif (lymphocytes T) afin de déclencher une réponse spécifique. Ces cellules sont appelées ainsi « Cellules Présentatrices d'Antigène » (APC). Ces cellules seront plus largement décrites dans la partie 1.4.



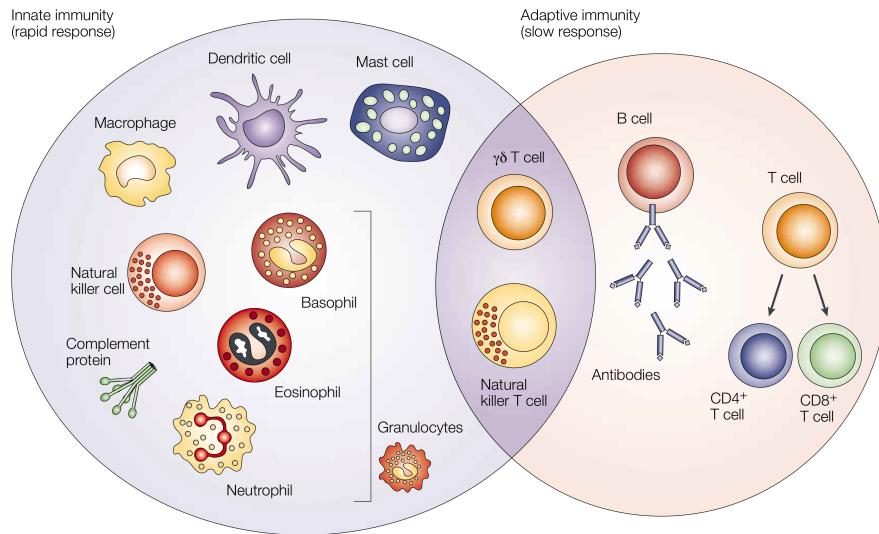


Figure 1.1 – Schéma représentant les principaux constituants cellulaires et solubles de l'immunité innée (cercle bleu) et de l'immunité adaptative (cercle rouge).[30]

Une autre partie importante du système immunitaire inné est le système du complément. Celui-ci est constitué de 35 protéines solubles du sérum dont 12 sont directement impliquées dans des mécanismes d'élimination des pathogènes, les autres permettant de réguler finement l'activité des premières. Ce système peut être activé par trois voies biochimiques différentes (Figure 1.2) :

- La voie classique passe par la reconnaissance d'une des protéines du complément (C1q) avec un complexe anticorps-antigène (reconnaissance par la partie Fc des IgG ou IgM) pour activer la cascade d'évènement conduisant à la dégradation du pathogène par la formation de pores.
- La voie alterne est activée directement par la reconnaissance de la protéine C3b avec des molécules à la surface des bactéries gram<sup>+</sup> et gram<sup>-</sup> entre autres, comme le LPS (lipopolysaccharide).
- La voie lectine, quant à elle passe par la fixation de la MBP (« Mannose Binding Protein ») sur la surface du pathogène. L'activation du système du complément par la MBP rejoint la voie classique par la suite.

Le système du complément a été défini comme appartenant au système immunitaire inné par les deux dernières voies d'activation qui n'impliquent aucune molécule du système immunitaire adaptatif.

Pendant longtemps les biologistes pensaient que les deux types d'immunités étaient indépendants, mais les cellules dendritiques et les macrophages sont pourtant essentiels aussi bien dans la dégradation des pathogènes (immunité innée) que dans l'activation des lymphocytes par la présentation de fragments des pathogènes (immunité adaptative). Ces

### 1.3. SYSTÈME IMMUNITAIRE ADAPTATIF

deux types cellulaires se situent donc à l'interface entre l'immunité innée et l'immunité adaptative. Ainsi, bien que l'on continue de décrire le système immunitaire selon ces deux composantes, il existe en réalité un continuum entre les deux.

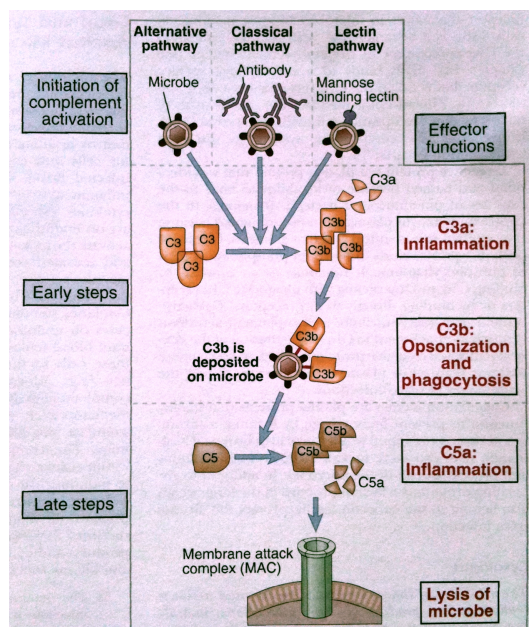


Figure 1.2 – Les voies du système du complément. (cellular and molecular immunology, fifth edition)

## 1.3 Système immunitaire adaptatif

Probablement, la première description d'une immunité acquise provient des récits de Thucydide -430 av JC. En effet, dans le livre II « Histoire de la guerre du Péloponnèse », il relate l'épidémie de peste à Athènes. Dans son récit, il explique que seules les personnes ayant guéri de la peste étaient capables de soigner les malades. C'était la première référence à l'acquisition d'une immunité par l'organisme.

Comme évoqué précédemment, la réponse de l'immunité adaptative est spécifique à une infection donnée mais nécessite un temps de latence. Cette immunité possède une caractéristique importante, la mémoire. En effet, une seconde infection d'un pathogène entraîne une réponse spécifique plus rapide. Les cellules de ce système immunitaire nécessitent d'être reconnues par des APCs qui permettent de sélectionner et d'amplifier les cellules possédant les récepteurs ou les anticorps spécifiques de la surface du pathogène. Dans cette partie, je développerai les différentes réponses à une activation par des APCs des lymphocytes. L'activation des lymphocytes sera présentée dans la partie 1.4.

Ce système immunitaire est constitué de deux principaux types de cellules : les lymphocytes T (T pour thymus) et les lymphocytes B (B pour « bone marrow »). Ces deux

types cellulaires ont longtemps (et encore maintenant) décrit deux mécanismes distincts : l'immunité à médiation cellulaire pour les lymphocytes T et l'immunité à médiation humorale pour les lymphocytes B (Figure 1.3). Ces cellules sont produites par la moelle osseuse mais diffèrent dans leur période de maturation. En effet, les lymphocytes B continuent leur maturation dans la moelle osseuse alors que les lymphocytes T, nouvellement produits, migrent dans le thymus pour la finir.

La deuxième différence majeure de ces deux populations de lymphocytes est le type de récepteur de l'antigène présent à la surface de ces cellules. Effectivement, les lymphocytes B matures possèdent des anticorps membranaires (BCR, « B cell receptor »). Au sein de cette population, chaque lymphocyte se différencie par un anticorps membranaire différent spécifique d'un antigène donné. Suite à la reconnaissance de l'antigène, le lymphocyte possédant l'anticorps correspondant se multiplie et sécrète un grand nombre d'anticorps solubles pour répondre à la menace.

Contrairement aux cellules B, les lymphocytes T présentent à leur surface des protéines appelées récepteur des cellules T (TCR, « T cell receptor »), mais à l'inverse des anticorps des cellules B, capables de reconnaître un antigène seul, les TCRs ont besoin que les antigènes soient présentés par une cellule présentatrice d'antigène (APC). Les APCs utilisent une protéine dénommée molécules du complexe majeur d'histocompatibilité (CMH) pour présenter les antigènes. Il existe deux classes de CMH qui diffèrent par leurs représentations à la surface des cellules mais aussi par l'interaction avec les deux sous-populations principales de lymphocytes T.

Le complexe majeur d'histocompatibilité de classe I est présent à la surface de pratiquement toutes les cellules et permet d'activer les lymphocytes T cytotoxiques ( $T_C$ ). Alors que le CMH de classe II est présent principalement sur les cellules présentatrices d'antigènes et permet d'activer les lymphocytes T auxiliaires (appelés  $T_H$  pour « T lymphocyte helper »). Ces deux sous-populations sont produites après la différenciation d'un lymphocyte T naïf. La différence de reconnaissance des TCR, des  $T_H$  et  $T_C$ , pour les CMH de classe I et II est due à la présence d'une glycoprotéine membranaire associée au TCR, appelée CD4 et CD8 respectivement. Ces glycoprotéines favorisent l'interaction des TCR avec le complexe CMH/antigène en reconnaissant la classe du CMH. Ainsi le marqueur CD4 est requis pour la reconnaissance du CMH de classe II alors que CD8 l'est pour les CMH de classe I (Figure 1.3).

## 1.4. LES CELLULES PRÉSENTATRICES D'ANTIGÈNES (APC)

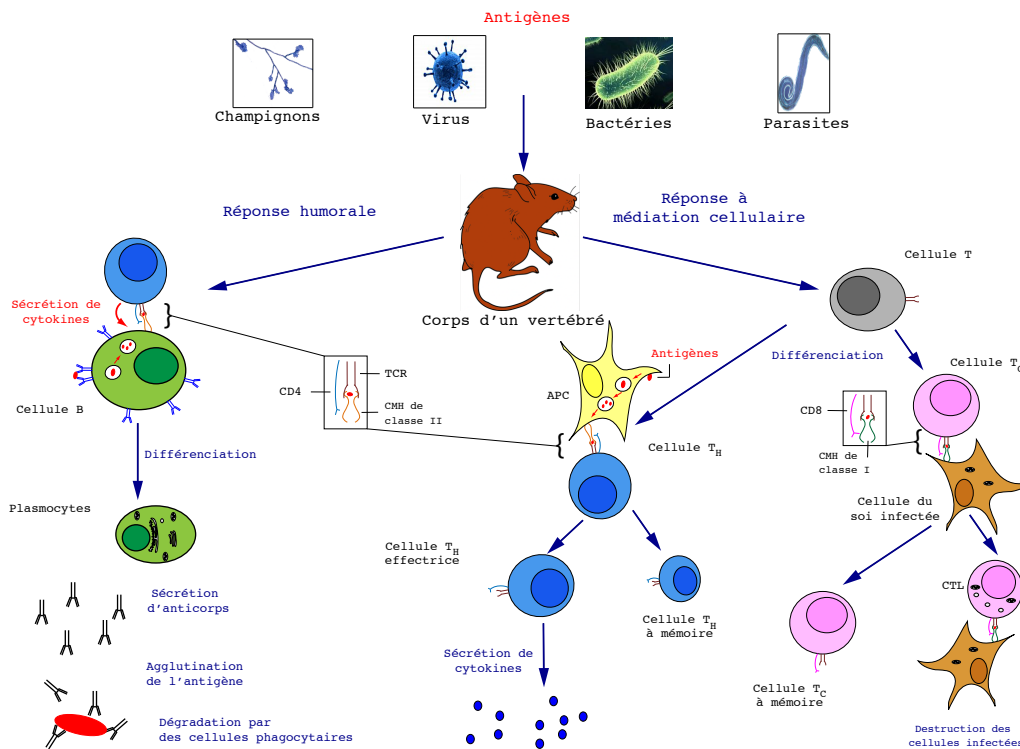


Figure 1.3 – Vue d'ensemble des branches humorale et à médiation cellulaire du système immunitaire. (adapté à partir de « Immunologie » de Janis Kuby).

La capacité des  $T_C$  à reconnaître la classe I de CMH, leur permet d'identifier et de détruire les cellules du soi altérées (cellules infectées par un virus par exemple) alors que les  $T_H$  activés par les APCs vont sécréter des cytokines qui vont permettre de stimuler la différenciation des cellules B, capables de reconnaître l'antigène, en plasmocytes sécréteurs d'anticorps solubles. Effectivement, les lymphocytes B matures vont se différencier en cellules B à mémoire et en plasmocytes après avoir reconnu l'antigène par ces anticorps membranaires et par la présence de cytokines. Les lymphocytes B font parties des cellules présentatrices d'antigènes car la reconnaissance de l'antigène par l'anticorps membranaire conduit à l'internalisation de l'antigène et sa présentation par le CMH de classe II au  $T_H$ .

## 1.4 Les Cellules Présentatrices d'Antigènes (APC)

Les cellules présentatrices d'antigènes sont des cellules expertes dans la reconnaissance d'un pathogène par la présence d'un large spectre de récepteur à leur surface, et dans l'internalisation de ceux-ci dans des compartiments cellulaires capables de les dégrader. A ce titre les macrophages et les cellules dendritiques font partie de cette famille de cellule. Ces cellules sont capables de reconnaître toutes sortes de molécules du « non-soi » par la présence d'un ensemble de motifs moléculaires spécifiques des pathogènes et

absents de l'hôte, regroupé sous le terme de « Pathogen Associated Molecular Pattern » (PAMP). Ces PAMPs sont reconnus grâce à de multiples familles de récepteurs spécifiques appelées « pattern recognition receptor » (PRR). A la suite de cette reconnaissance, ces cellules internalisent dans des phagosomes les pathogènes pour les dégrader et présenter des peptides, par les molécules du CMH (de classe I et II), aux lymphocytes T afin de déclencher une réponse immunitaire spécifique.

La reconnaissance de pathogène par les PRRs déclenche donc l'internalisation de celui-ci dans un compartiment cellulaire dénommé phagosome. Le phagosome est formé par l'invagination de la membrane plasmique. Par la suite ces phagosomes évoluent en fusionnant avec des lysosomes afin de dégrader le microbe. Pour ce faire, la cellule présentatrice d'antigène utilise les enzymes présentes dans les lysosomes et toute la machinerie d'acidification des compartiments endocytiques [15]. On retrouve de nombreuses enzymes lytiques dans ces compartiments qui permettent la digestion aussi bien de polysaccharides, d'ADN, d'ARN que de protéines. La dégradation de l'antigène du « non-soi » est essentielle pour présenter les fragments aux lymphocytes T par le CMH. Un autre mécanisme important pour la microbicidie et la dégradation est l'activation de la NADPH oxydase qui transforme le dioxygène en espèce réactive de l'oxygène (ROS) toxique pour les micro-organismes. Mais la capacité des phagosomes à dégrader les molécules n'est pas similaire pour les macrophages et les cellules dendritiques [87]. En effet, les macrophages possèdent une activité NADPH oxydase supérieure à celle des cellules dendritiques conduisant à une dégradation plus poussée du pathogène [87]. La principale différence entre les phagosomes des macrophages et ceux des cellules dendritiques est l'évolution du pH. Effectivement, le pH des phagosomes des cellules dendritiques ne bouge presque pas dans le temps alors que celui des macrophages s'acidifie fortement avec le temps, cette acidification permet aux hydrolases acides d'exercer leur activité dans des conditions optimum (Figure 1.4).

Cette régulation de la digestion de micro-organismes permet aux APCs de présenter des fragments différents de peptides antigéniques aux lymphocytes. Ainsi, une dégradation plus mesurée dans les cellules dendritiques permet l'obtention de peptides plus importants et une meilleure présentation pour la sélection des lymphocytes T spécifiques.

Après la digestion des micro-organismes présents dans les phagosomes, ceux-ci fusionnent avec la membrane plasmique (mécanisme d'exocytose) pour libérer les débris générés et présenter à la surface cellulaire des peptides couplés aux molécules du CMH.

Pendant toute cette phase de digestion d'agents infectieux dans les phagosomes, les APCs migrent de leur lieu de rencontre avec l'agent infectieux vers des organes lymphoïdes secondaires afin d'activer les lymphocytes.

#### 1.4. LES CELLULES PRÉSENTATRICES D'ANTIGÈNES (APC)

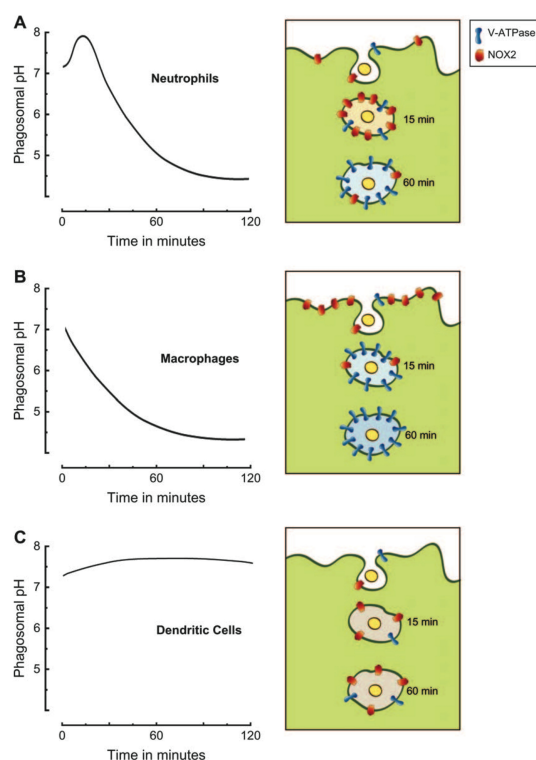


Figure 1.4 – Représentation du pH phagosomal en fonction du temps pour les neutrophiles (A), les macrophages (B) et pour les cellules dendritiques (C). V-ATPase est une pompe à proton permettant l'acidification du phagosome. NOX2 est une isoforme de la NADPH oxydase.[87]

Dans la partie suivante, nous allons nous intéresser aux cellules dendritiques dans lesquelles la protéine étudiée dans ce travail de thèse est exprimée.

## Chapitre 2

### Les cellules dendritiques (DC)

Tout d'abord, dans un souci de simplification, j'ai décrit auparavant les cellules dendritiques comme étant présentes dans les tissus, or on retrouve aussi des cellules dendritiques circulantes dans le sang. La plus large population de cellules dendritiques est néanmoins tissulaire.

La première barrière physique rencontrée par un pathogène à l'interface entre le milieu extérieur et notre organisme est la peau et les muqueuses. Cette barrière est composée de deux couches différentes, la première est l'épiderme qui constitue la partie en contact avec le milieu extérieur et la seconde est le derme (Figure 2.1).

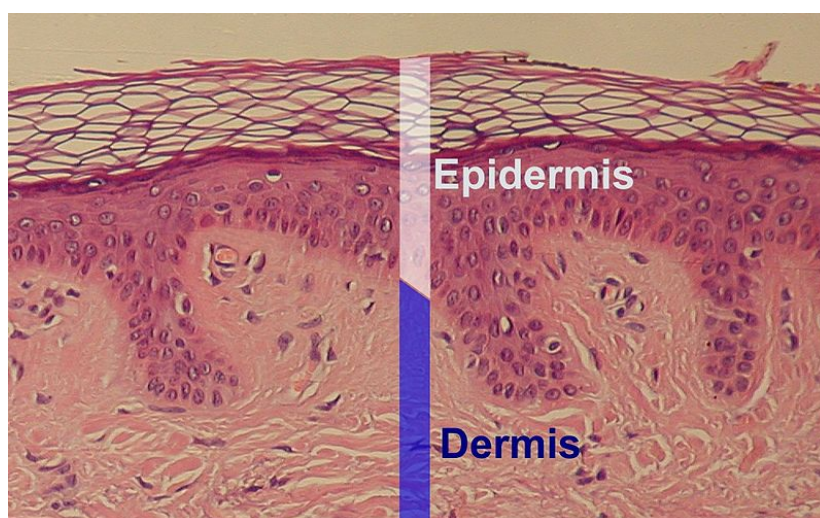


Figure 2.1 – Micrographie de peau indiquant les deux couches principales de cellules. [83]

Cette barrière empêche un grand nombre de pathogènes d'infecter le corps humain mais celle-ci peut être franchie lors de lésions ou de plaies, par exemple. Certaines cellules du système immunitaire, présentes dans cet environnement, peuvent détecter l'intrusion de pathogènes. Ces cellules jouent un rôle de « sentinelles » capables d'alerter le système



immunitaire et contribuent une large famille cellulaire appelée cellule dendritique (DC). Parmi celles-ci, on retrouve notamment deux types de cellules dendritiques :

- Les cellules de Langerhans résidant dans l'épiderme
- Les cellules dendritiques dermiques présentes dans le derme comme leur nom l'indique.

Suite à la reconnaissance d'un agent pathogène, ces cellules sont activées et migrent jusqu'aux organes lymphoïdes secondaires afin de stimuler les lymphocytes T spécifiques (Figure 2.2).

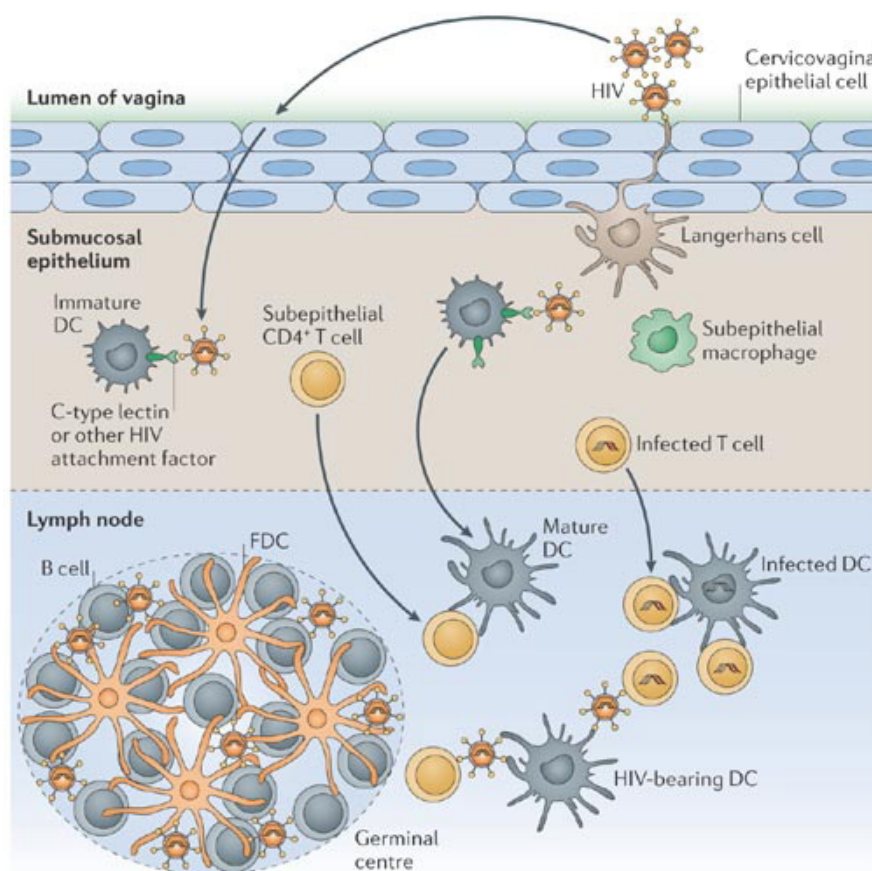


Figure 2.2 – Localisation et migration de DCs dans le cas d'une infection par le VIH.[115]

Les cellules dendritiques sont définies selon trois critères : ces cellules sont capables de présenter des antigènes via le CMH de classe II, d'activer les lymphocytes T naïfs et ont une morphologie dendritique [93]. Cette morphologie dendritique est caractérisée par la présence d'excroissances ressemblant aux axones des neurones (Figure 2.3). Ces excroissances sont appelées dendrites.



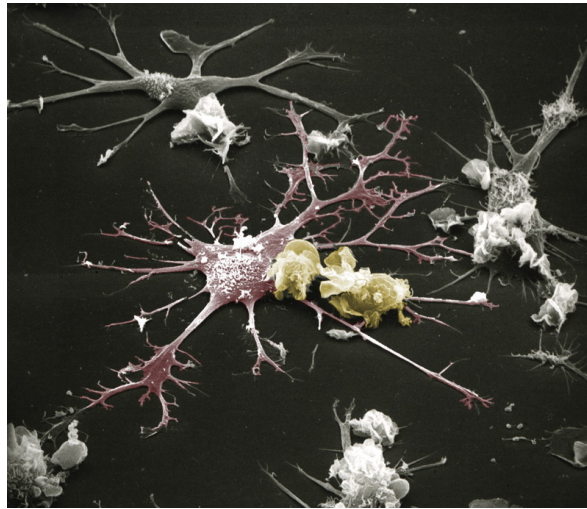


Figure 2.3 – Image de cellules dendritiques. <http://www.uni-ulm.de/expane/images/dendritic.jpg>

## 2.1 Les différentes cellules dendritiques et leurs rôles

Les cellules dendritiques possèdent des fonctions essentielles qui vont de l'initiation et de la régulation de l'immunité innée et adaptative à l'établissement de mécanismes de tolérance centrale dans le thymus [77]. On les retrouve aussi impliquées dans des cancers [74] et des phénomènes de régulation d'hypersensibilité de contact (allergie) [75] montrant d'autant plus leur importance dans la régulation de la réponse immunitaire. Les DCs constituent une population hétérogène de cellules d'origine hématopoïétique donnant deux principaux progéniteurs. Un progéniteur lymphoïde qui va se différencier en cellule dendritique lymphoïde (plus généralement appelée cellule dendritique plasmacytoïde, l'origine lymphoïde est encore controversée). Un progéniteur myéloïde qui va se différencier pour donner deux précurseurs sanguins, des monocytes ( $CD14^+ CD11c^+ CD1^-$ ) et des cellules  $CD14^- CD11c^+ CD1^+$  qui vont eux même se différencier pour donner des cellules dendritiques interstitielles (ou des macrophages, même précurseur) et des cellules de Langerhans (ou macrophages de « Langerhans ») (Figure 2.4) [4]. La différenciation des précurseurs en différents types de DCs est dépendante de l'environnement de cytokines présents comme les interleukines (IL).

Les cellules dendritiques plasmacytoïdes ont tout d'abord été décrites comme « des cellules T plasmacytoïdes » dues à leur localisation dans les zones T des ganglions lymphatiques dans un contexte particulier de lymphadénite non spécifique et de leur morphologie proche de celle des plasmocytes [109]. Par la suite elles ont été définies comme des monocytes plasmacytoïdes par leur expression de marqueurs de la lignée monocyte-macrophage ( $CD31$ ,  $CD36$ ...) et ce n'est qu'à la fin des années 90 que plusieurs équipes de recherche

## 2.1. LES DIFFÉRENTES CELLULES DENDRITIQUES ET LEURS RÔLES

ont démontré que ces cellules se différenciaient *in vitro* en cellules dendritiques [49] et furent dénommées « cellules dendritiques plasmacytoïdes » (pDC).

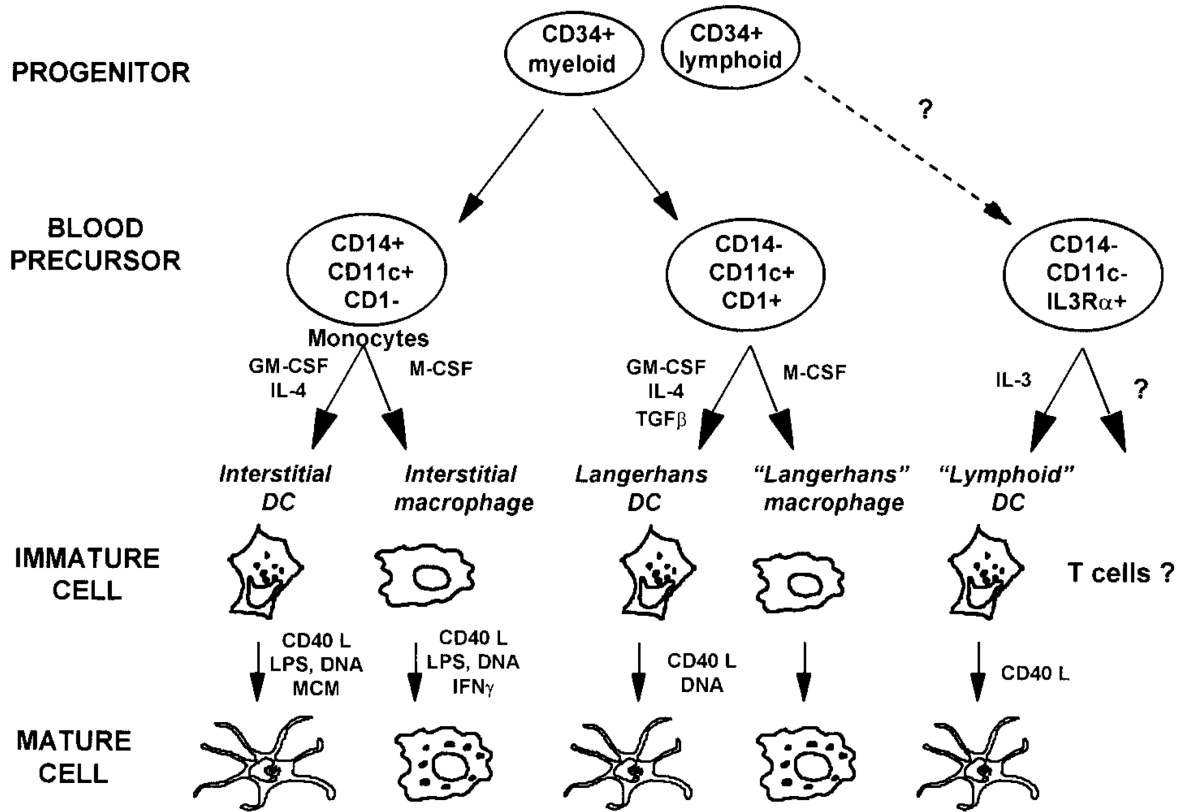


Figure 2.4 – Représentation schématique des sous-populations de cellules dendritiques et de leurs précurseurs.[4]

Ces cellules dendritiques sont particulières par leur capacité sécrétrice importante d'interféron (IFN) de type 1 en réponse à une infection virale [13][91]. Elles ont été alors appelées cellules naturelles productrices d'IFN (IPC), et se caractérisent par un réticulum endoplasmique abondant ainsi qu'un appareil de Golgi développé [49]. Les pDCs circulantes expriment des marqueurs qui leur sont propres comme BDCA-2 et BDCA-4 (Blood Dendritic Cell Antigen) [34].

Les DCs nous intéressant particulièrement sont les cellules dendritiques myéloïdes et plus précisément les cellules de Langerhans qui seront développées dans une autre partie (2.4). Ces cellules jouent un rôle essentiel de sentinelle du système immunitaire par leur localisation au niveau des entrées de l'organisme (peau, muqueuse)[3]. Elles possèdent un large spectre de PRR, comme les TLR ou encore les récepteurs lectines de type-C, leur permettant de reconnaître tout type de motif moléculaire conservé à la surface d'organisme du « non-soi » (PAMPs). Mais elles peuvent aussi être impliquées dans des maladies graves

en étant les cellules cibles d'un pathogène comme dans le cas de l'infection par le virus de la dengue [116] ou encore être responsables de cancer [74].

## 2.2 La reconnaissance du « non-soi » par les cellules dendritiques : les « Toll-Like Receptors » et les lectines de type-C

Les cellules dendritiques possèdent un large spectre de récepteurs pour reconnaître aussi bien des motifs saccharidiques nucléotidiques, peptidiques ou lipidiques à la surface de pathogènes. Ces agents étrangers possèdent à leur surface des molécules conservées de différentes natures appelées PAMPs (« Pathogen-Associated Molecular Pattern »). Parmi les PAMPs, on trouve, entre autre, le lipopolysaccharide (LPS) présent à la surface des bactéries, la flagellin, des acides nucléiques associés aux virus comme les ARN double brins (dsRNA) ou des glycoprotéines par exemple. Pour reconnaître ce nombre élevé de PAMPs, les cellules ont développé un grand nombre de récepteurs comme les TLRs (« Toll-Like Receptor ») ou encore les CLRs (« C-type Lectin Receptor »). Le chapitre 4 sera destiné aux lectines donc je vais brièvement évoquer les TLRs et les lectines de type-C.

### 2.2.1 Les récepteurs de type Toll (TLR)

Chez l'humain, il existe une dizaine de TLRs avec une spécificité de reconnaissance différente (Figure 2.5). Ces récepteurs sont capables aussi bien de reconnaître les bactéries que les virus. Après la reconnaissance de l'agent pathogène, ces récepteurs déclenchent des cascades protéiques (comme les MAP kinases) menant, pour la plupart, à l'activation du facteur de transcription NF $\kappa$ B [5, 29]. Celui-ci permet d'activer la transcription des gènes codant pour des cytokines qui vont eux-mêmes conduire à différents processus tel que l'apoptose de la cellule, l'activation de cellules immunitaires ou encore la destruction de l'agent pathogène.

Les différentes populations de cellules dendritiques n'expriment pas à leur surface les mêmes TLRs (Tableau 2.1). Cette distribution différente des TLRs à leurs surfaces induit une spécificité différente dans la reconnaissance d'agent pathogène.

## 2.2. LA RECONNAISSANCE DU « NON-SOI » PAR LES CELLULES DENDRITIQUES : LES « TOLL-LIKE RECEPTORS » ET LES LECTINES DE TYPE-C

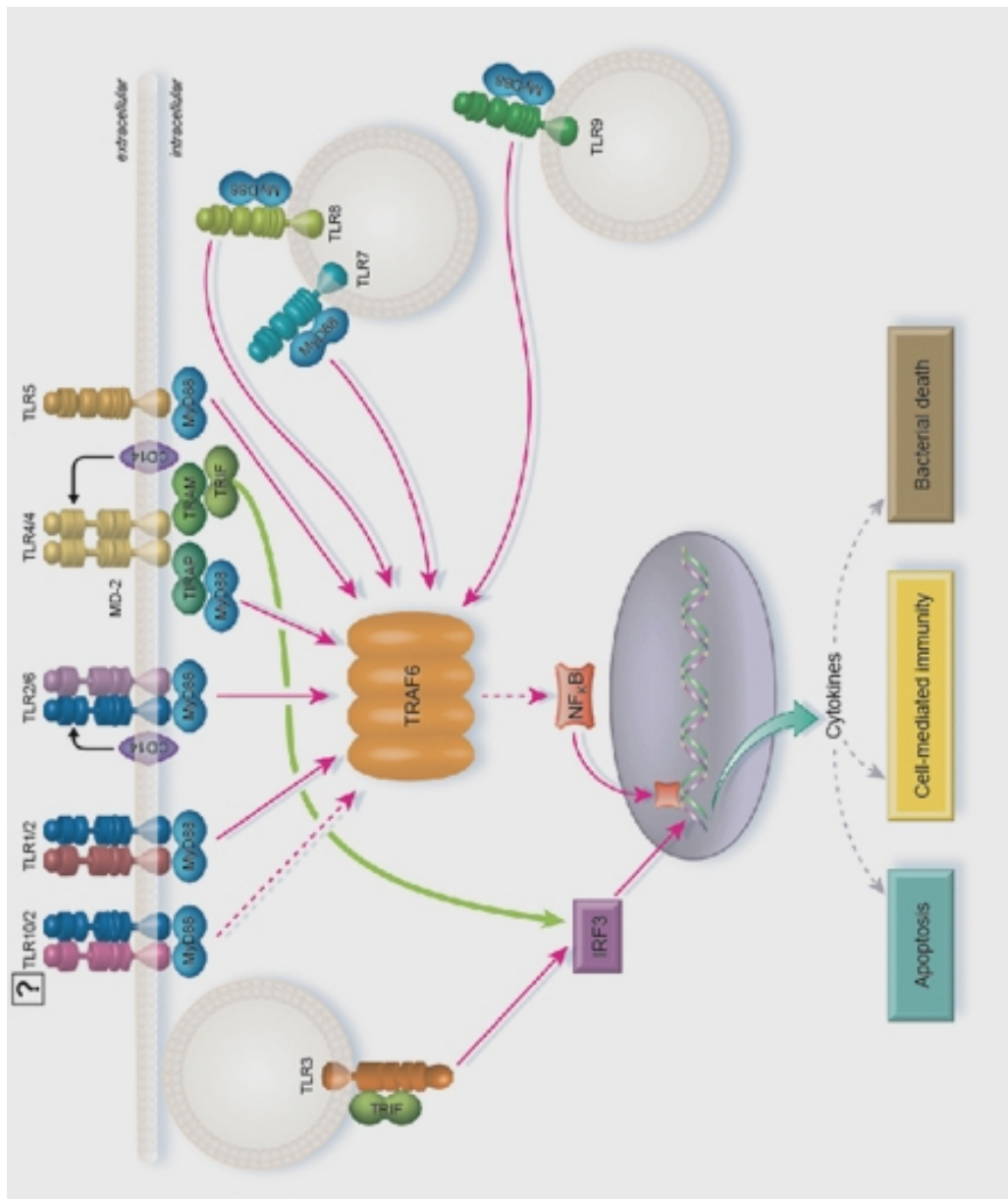


Figure 2.5 – Représentation des différents TLRs, leurs ligands et les voies de signalisation impliquées. dsRNA : ARN double brin. ssRNA : ARN simple brin. IRF3 : facteur de transcription de l'interféron. NF $\kappa$ B : facteur de transcription. Image réalisée par l'entreprise ABD SEROTEC [http://www.abdserotec.com/minireview\\_toll\\_like\\_receptors\\_tlrs\\_-1060.html](http://www.abdserotec.com/minireview_toll_like_receptors_tlrs_-1060.html)

DC subtype	TLR	Expression	Effects of activation
moDC	1	+	See TLR2
	2	++	Increased IL-6, IL-8, IL-10, IL-12, TNF $\alpha$ Low IFN $\beta$ response and no IFN $\alpha$ response
	3	+	Specific IFN $\beta$ mRNA upregulation (not IFN $\alpha$ )
	4	++	Upregulation of CD80, CD86, CD83, CCR7 Secretion of IFN $\beta$ , IFN $\gamma$ , IL-1 $\beta$ , IL-12p70, IL-13, IP-10 Decreased endocytic capacity
	5	+	Upregulation of CD80, CD86, CD83, CCR7 Secretion of IFN $\gamma$ , IL-1 $\beta$ , TNF, IL-8, IL-12p40 (not IL-12p70), IL-13 Decreased endocytic capacity
	6	$\pm$	See TLR2
	7	$\pm$	–
	8	+	Increased TNF $\alpha$ , IL-8, IL-12p40, MCP-1, CCL2, CCL3, CCL4, CCL5
	9	–	
	10	–	
mDC	1	+	See TLR2
	2	++	Upregulation of CCR7, IL-6, IL-10, IL-12p70, TNF $\alpha$ , no IFN $\alpha$
	3	++	IFN $\alpha$ (intermediate), IL-12p70 (high) No TNF $\alpha$ or IL-6
	4	+	Upregulation of CD80, CD86, CD83, CD40, CCR7 Secretion or upregulation of CCR7, IL-6, IL-8, IL-10, IL-12p70 No IFN $\alpha$ response
	5	+	Upregulation of CD80, CD86, CD83, CCR7 Secretion of TNF and IL-8 Upregulation of CCR7
	6	+	See TLR2
	7	+	Upregulation of CD40, CD80 and CD86 Secretion of IL-12p70 No IFN $\alpha$ response
	8	$\pm$	See TLR7
	9	–	
	10	+	No ligand known
pDC	1	$\pm$	
	2	–	
	3	–	
	4	–	
	5	–	
	6	–	
	7	++	Upregulation of CD40, CD80, CD86, CCR7 Very high IFN $\alpha$ response No IL-12p70 response
	8	–	
	9	+++	Upregulation of CD40, CD80, CD86, CD83, HLA-DR, CCR7 Upregulation of IFN $\alpha$ (very high), IFN $\beta$ (lower), IL-6, TNF $\alpha$ (low), IL-8, IP-10 No IL-10 secretion
	10	+	No ligand known

Tableau 2.1 – Expression des TLRs dans les différents types de cellules dendritiques. moDC : cellule dendritique dérivée de monocyte. mDC : cellule dendritique myéloïde. pDC : cellule dendritique plasmacytoïde.[89]

### 2.2.2 Les lectines de type-C

Les lectines de type-C représentent une famille importante de PRR dans les cellules dendritiques. Un certain nombre est largement exprimé comme par exemple la Langérine, DC-SIGN, la MGL (« Macrophage Galactose Lectin »), et les Dectines 1 et 2 (Tableau 2.2). Ces lectines permettent de reconnaître des motifs saccharidiques présent à la surface de nombreux pathogènes tels que les bactérie avec le LPS, les virus avec leurs glycoprotéines associées . . . , et d’internaliser ces molécules du « non-soi » dans des phagosomes pour leur dégradation. Ce processus active la migration et la maturation des DCs qui finis par la présentation d’antigène aux lymphocytes T.

C-type lectin	Type	Amino acids	Chromosomal location	Production	Ligand	Function	Key antibodies
MMR (CD206)	I	1456	10p13	DCs, LCs, Mo, Mφ, DMECs	Mannose, fucose, sLeX	Antigen uptake <sup>5</sup>	MG38, anti-human <sup>36,47</sup>
DEC-205 (CD205)	I	1722	2q24	DCs, LCs, high on actDCs, thymic ECs	?	Antigen uptake <sup>54</sup>	
Dectin 1	II	247	12p13	DCs, LCs	β-Glucan <sup>28</sup>	T-cell interaction <sup>27</sup>	
Dectin 2	II	209	2p13	DCs, LCs	?	Antigen uptake <sup>23</sup>	
Langerin (CD207)	II	328	2p13	LCs	?	Formation of Birbeck granules <sup>35</sup>	DCGM4, anti-human
DC-SIGN (CD209)	II	404	19p13	DCs	HIV-1 (gp120), SIV, mannan, ICAM-2, ICAM-3	T-cell interaction <sup>15</sup> , HIV-1 pathology <sup>17</sup> , migration <sup>34</sup> , antigen uptake	AZN-D1, anti-human
BDCA-2 (REF. 3)	II	?	?	Plasmacytoid DCs	?	Antigen uptake?	AC-144, anti-human
DCIR (REF. 22) (LLIR) <sup>25</sup>	II	237	12p13	DCs, Mo, Mφ, PMN, B	?	?	(REF. 25)
DLEC <sup>26</sup>		231	12p13				
CLEC-1 (REF. 24)	II	280 / 229	12	DCs	?	?	

Abbreviations: actDCs, activated dendritic cells; B, B cells; DMECs, dermal microvascular endothelial cells; Mo, monocytes; Mφ, macrophages; PMN, polymorphic nuclear cells; sLeX, sialyl Lewis X; thymic ECs, thymic endothelial cells.

Tableau 2.2 – Liste succincte des lectines de type-C exprimées par les cellules dendritiques. [41]

## 2.3 La maturation des cellules dendritiques

A partir de la rencontre avec un élément du « non-soi », les DCs modifient radicalement les protéines à leur surface afin de se détacher de la matrice extracellulaire pour migrer par le réseau lymphatique vers les organes lymphoïdes secondaires (Figure 2.2). Parallèlement, les DCs activent les voies d’apprêtement des antigènes et déclenchent la maturation du phagosome (Figure 1.4, Figure 2.6) [4].

En effet après la capture du pathogène, les récepteurs aux chimiokines ainsi que les PRRs ont leur expression diminuée par la cellule. Par contre celle-ci augmente l’expression des molécules du CMH de classe II, afin de présenter les antigènes aux lymphocytes, ainsi que la protéine CCR7 qui permet la domiciliation de la cellule au niveau des organes

lymphoïdes secondaires. De plus, la morphologie de ces cellules est modifiée, on remarque l'apparition de dendrites.

## Différentiation et maturation des cellules dendritiques

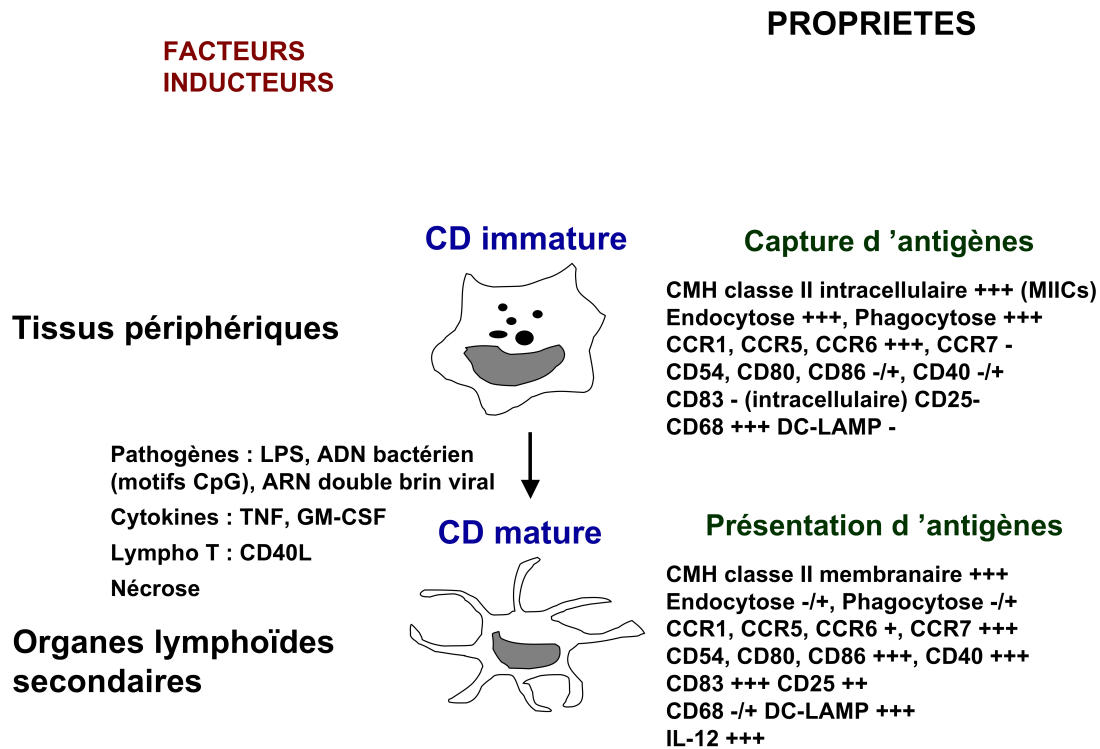


Figure 2.6 – Maturation des cellules dendritiques. A gauche sont montrés les différents facteurs déclenchant la maturation. A droite sont représentées les principales propriétés de chaque étape de la maturation/différentiation.[4]. traduit par [www.sante.univ-nantes.fr](http://www.sante.univ-nantes.fr).

## 2.4 Les cellules de Langerhans (LCs)

Les cellules de Langerhans ont été découvertes par Paul Langerhans en 1867 pendant ses études de médecine à l'Institut de Pathologie de Berlin [65]. Ce ne fut que bien plus tard que ces cellules ont été décrites comme des cellules dendritiques qui jouent un rôle dans la médiation cellulaire de l'immunité. A partir des années 60, Michael Stanley Clive Birbeck et d'autres chercheurs se sont intéressés à étudier les cellules de Langerhans par microscopie électronique et ils ont découvert que ces cellules possèdent un organe cytoplasmique avec une structure particulière en « raquette de tennis » (Figure 2.7) [8, 114]. Cet organe spécifique de ces cellules fût appelé le granule de Birbeck.

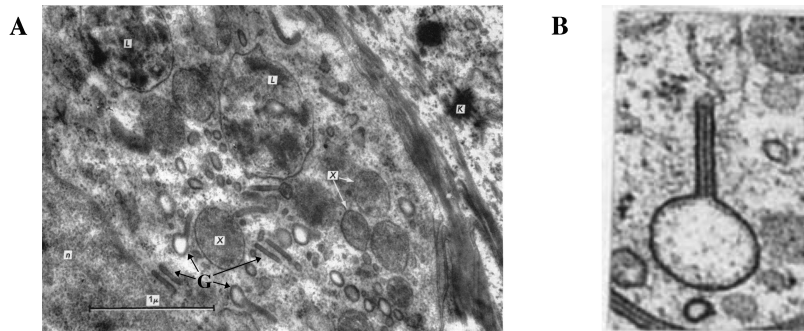


Figure 2.7 – Micrographie d’une cellule de Langerhans (A) et d’un agrandissement sur un granule de Birbeck (B). G : granule de Birbeck, K : keratinocyte adjacent, L : lysosome, n : noyau, X : organite rond avec une structure interne granulaire.[8]

Les cellules de Langerhans ont été initialement définies comme un sous-type de cellule dendritique unique pour deux raisons principales :

- La présence du granule de Birbeck.
- Par la localisation tissulaire de ces cellules, au niveaux des épidermes.

### 2.4.1 La localisation des cellules de Langerhans

Les cellules de Langerhans sont présentes dans les couches supérieures des épithélia et elles sont les principales APCs présentes dans ce tissu. Cette localisation fait de ces cellules les premières cellules dendritiques à être en contact avec des agents pathogènes et accroît le rôle de sentinelle du système immunitaire.

### 2.4.2 Les différentes fonctions des cellules de Langerhans

Comme toutes cellules dendritiques, elles sont capables d’activer les lymphocytes T naïfs afin de répondre efficacement à une menace qui se prolonge dans le temps. Mais les cellules de Langerhans sont aussi impliquées dans différents phénomènes tels que l’induction et le contrôle de la réponse à l’hypersensitivité de contact (ou allergie) [59] ou encore dans une maladie rare, le cancer des cellules de Langerhans (LCS pour Langerhans cell sarcoma) [66, 74].

Mais le rôle principal de ces cellules reste la reconnaissance de molécules du « non-soi ». Elles sont capables de reconnaître toutes sortes de champignons, tels que *Malassezia furfur* et des espèces de *Candida* et de *Saccharomyces* [99, 23], ainsi que des mycobactéries, avec en particulier *Mycobacterium leprae*[55], ou encore des virus comme le virus d’immunodéficience humaine et l’herpex simplex 2 [22, 21, 24].



### 2.4.3 Le granule de Birbeck

Le granule de Birbeck (BG) est un organite cytoplasmique spécifique des cellules de Langerhans. Dans les années 60, de nombreuses études de microscopie électronique ont été réalisées et des premières reconstructions de cette organite ont été effectuées [8, 114, 84]. Cette structure a été caractérisée de plusieurs façons comme « en forme de disque » puis « en forme de raquette de tennis » qui est encore utilisée actuellement (Figure 2.8) [84].

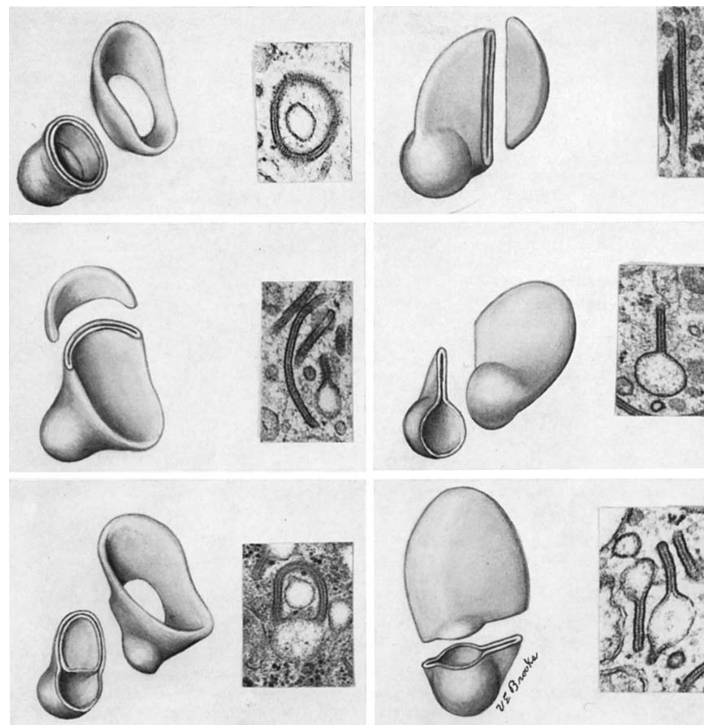


Figure 2.8 – Modèles de coupe de granule de Birbeck dans différents plans.[84]

Ce n'est que dans les années 2000 que fut identifiée la protéine responsable de la formation de ce granule unique [104, 100]. Tout d'abord, il fut montré que la Langérine, une lectine de type-C, est présente dans ces structures de repliement de membrane plasmique, mais aussi que son « activation » induit une rapide internalisation dans le granule de Birbeck [104]. Dans cette même étude, Valladeau et al mettent en évidence que la seule transfection du gène de cette lectine dans des fibroblastes murins suffit à la formation de structure de type sandwich de cytomembrane (CMS) proche de la structure des granules de Birbeck (Figure 2.9). Par cette observation, on peut en déduire que la formation de cet organite intracellulaire unique est majoritairement due à l'expression d'une seule protéine, la Langérine.

La capacité de la Langérine à reconnaître un pathogène et à conduire à son internalisation dans les BGs a été illustrée pour la première fois vis-à-vis du VIH. Ceci sera

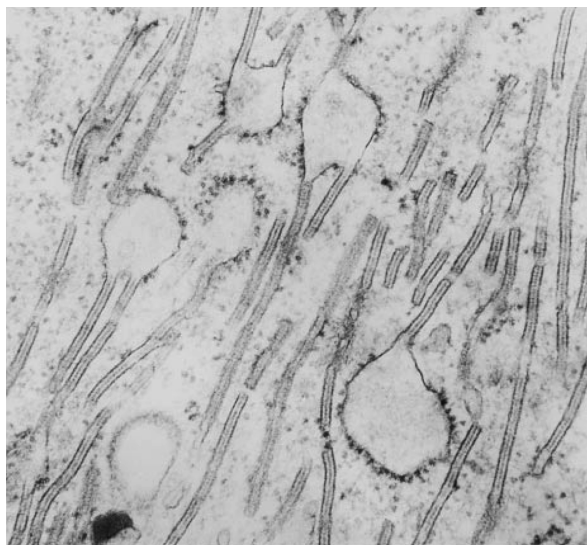


Figure 2.9 – L’expression de Langérine recombinante induit la formation de BG dans des fibroblastes murins.[104]

plus largement commenté dans le chapitre suivant mais a été à l’origine d’un regain d’intérêt pour cette protéine. Dans ce contexte, notre laboratoire s’est engagé dans des études de caractérisation biochimique structurale et fonctionnelle de cette protéine ainsi qu’une étude *in cellulo* au travers d’une collaboration. Ainsi par la transfection de gènes contenant différentes constructions de la Langérine dans des fibroblastes, et d’observations par microscopie électronique, nous avons pu étudier le rôle de la Langérine dans la formation des granules de Birbeck (BG) [100].

La seule fonction récemment identifiée du granule de Birbeck est un mécanisme de protection vis-à-vis de l’infection par *Mycobacterium leprae* [55], et par le virus d’immunodéficience acquise humain (VIH)[24]. Ce rôle sera développé dans le chapitre 4 consacré aux différents rôles, positif ou négatif, des cellules dendritiques face à l’infection par le VIH. A part ce rôle, il a été aussi montré que l’absence de la Langérine, et par conséquent de BG, n’affecte pas le développement de souris K.O (knock out) ni les fonctions de capture d’antigène, de migration et de maturation [62].

### 2.4.4 La reconnaissance de pathogène par les cellules de Langerhans

Les cellules de Langerhans expriment un répertoire de TLRs particulier qui permet à ces cellules de reconnaître aussi bien des bactéries Gram+ que des virus. Effectivement, les cellules de Langerhans expriment principalement les TLRs 1/2/3/5/6 et 10. Ces TLRs permettent donc de reconnaître le peptidoglycane et l’acide lipoteichoïque des

bactéries Gram+ (TLR1/2/6) ainsi que les ARN double brin présents chez certains virus (TLR3)[42].

Les cellules de Langerhans expriment à leur surface différents PRRs qui servent de marqueur de ces cellules. Parmi les PRRs exprimés, la Langérine (CD207) permet de discriminer ces cellules des autres cellules dendritiques par son fort taux d'expression. Son expression permet à ces cellules de reconnaître des champignons opportunistes et pathogène tels que *Candida Albicans* (Figure 2.10)[23]. Cette reconnaissance implique les propriétés « lectines » de la Langérine vis-à-vis des  $\beta$ -glucans de surface de ces champignons.

Enfin, la Langérine a été également impliquée dans la reconnaissance de virus de l'immunodéficience humaine (VIH) [24]. Ce point est détaillé dans le chapitre suivant et constitue une partie importante pour les bases du projet.

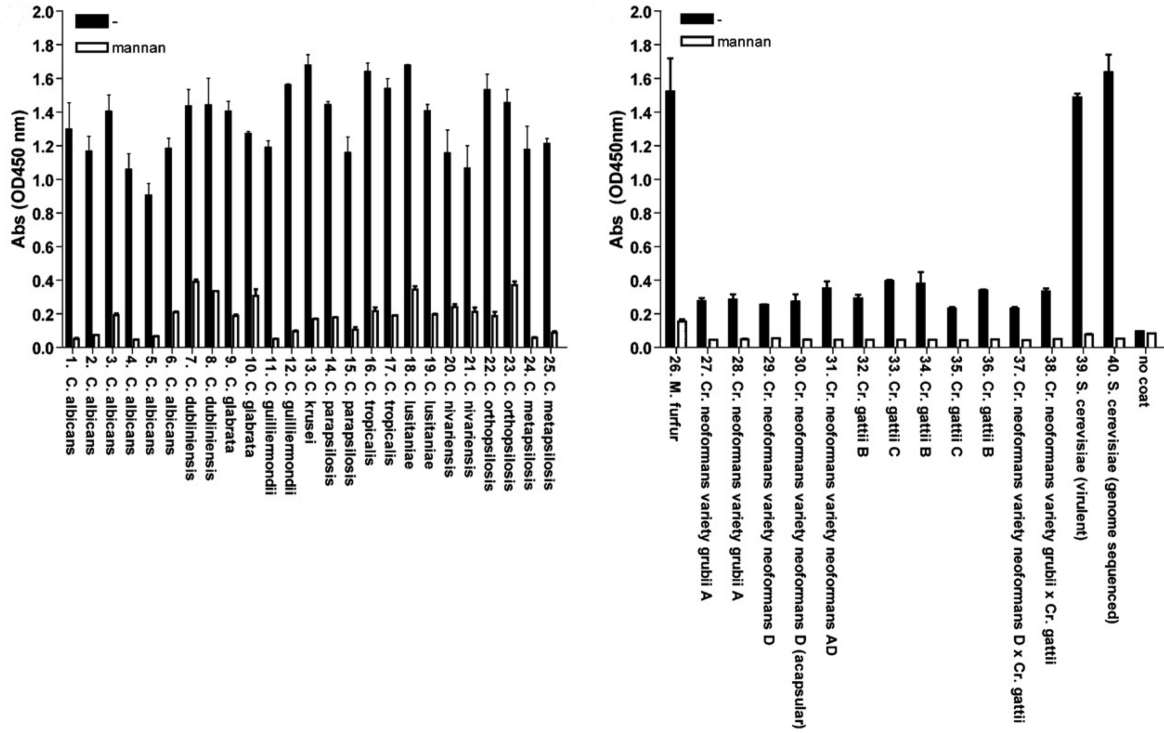


Figure 2.10 – Reconnaissance de champignons par la Langérine. [23]

---

## Chapitre 3

### Les cellules dendritiques et le VIH

Le VIH est un virus du genre lentivirus (du latin lenti, qui signifie lent) de la famille des rétrovirus et qui se caractérise par une longue période d'incubation. Le génome rétroviral est composé de deux copies identiques d'un ARN simple brin. Le virus est constitué d'une membrane lipidique dans laquelle est insérée la protéine gp191 (composée de gp41 et gp120), à l'intérieur de cette membrane, on trouve une matrice protéique (p17) puis une capside dans laquelle se trouve les deux ARN simple brin et des transcriptases inverses (Figure 3.1)[36]. Le VIH affecte notre système immunitaire par sa réplication dans ces cellules cibles, les lymphocytes T. Cette réplication intensive conduit à la lyse des cellules et la dissémination de nouveaux virus. La fragilisation du système immunitaire conduit à une immunodéficience grave.

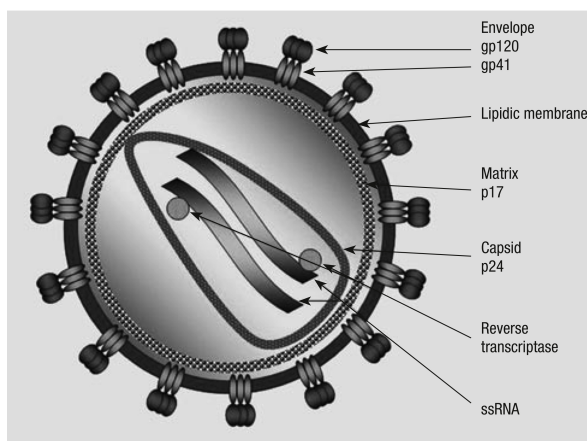


Figure 3.1 – Schéma d'une particule de VIH. [36]

La transmission sexuelle du VIH est la voie majeure d'infection. Mais pour cela, il faut passer la barrière des muqueuses. Il a été montré que des micro-lésions physiques, causées lors de relations sexuelles, ou de micro-abrasions résultants d'ulcère génital ou d'autres infections (syphilis, herpès génital ...) sont des facteurs augmentant les risques d'infection

[22, 54]. Au sein des muqueuses génitales, les premières cellules de l'immunité à même d'interagir et de capturer le virus sont les cellules de Langerhans et les cellules dendritiques dermiques. Ces deux populations de cellules dendritiques se comportent différemment lors de l'infection par le VIH. En effet, les cellules de Langerhans, contrairement aux DCs dermiques, possèdent un rôle protecteur en internalisant le virus dans les granules de Birbeck (BG) et cela grâce à la capacité de reconnaissance de la Langérine pour le VIH [24]. Par contre, les cellules dendritiques dermiques, n'exprimant pas la Langérine mais une autre lectine de type C (DC-SIGN) capable de reconnaître le VIH, sont détournées par le virus pour accéder aux organes lymphoïdes secondaires afin d'infecter des lymphocytes T, la cible du virus. La reconnaissance du virus par DC-SIGN active la migration des DCs dermiques. Pendant cette migration, le virus soit est exposé à la surface cellulaire par DC-SIGN afin d'infecter les lymphocytes T dans les organes lymphoïdes secondaires (trans-infection, Figure 3.2), soit d'infecter directement les DC dermiques conduisant à la multiplication du virus dans la cellule (cis-infection) [46]. Cette multiplication du virus dans la DC dermique, pendant la migration, permet d'exposer un grand nombre de virions dans les organes lymphoïdes secondaires pour infecter les lymphocytes T.

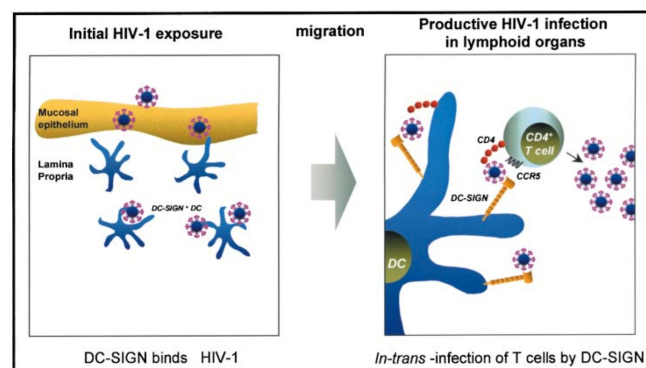


Figure 3.2 – Schéma d'un mode d'infection du VIH. [46]

Nous allons voir dans cette partie que l'expression de certains PRRs modifie la réponse aux virus par différentes sous-populations de DCs (cellules de Langerhans et les cellules dendritiques dermiques).

Une première partie est consacrée au détournement des cellules dendritiques dermiques par le VIH et une seconde partie présente un mécanisme de défense vis-à-vis du VIH par les cellules de Langerhans.

### 3.1 La trans-infection des lymphocytes T par les cellules dendritiques

Le concept de trans-infection consiste au recrutement d'une cellule par le virus afin d'infecter plus facilement la cellule cible. Nous verrons dans ce chapitre les différents mécanismes qui conduisent à cette trans-infection dans le cas du VIH.

Comme évoqué précédemment, le VIH utilise la faculté des cellules dendritiques dermiques à le reconnaître pour être dirigé au niveau des glandes lymphoïdes secondaires afin d'infecter les cellules cibles du VIH, les lymphocytes T [101]. Suite à la reconnaissance du virus par la lectine DC-SIGN, différents devenir du virus ont été décrits et documentés dans la littérature et résumés dans la figure 3.3.

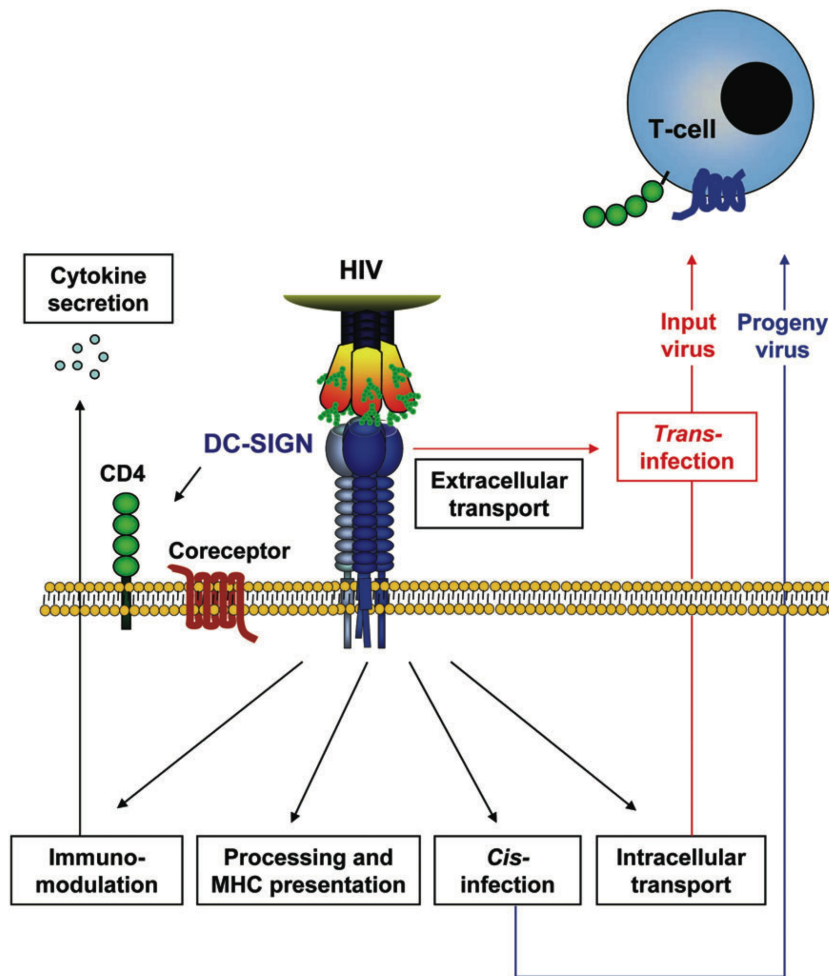


Figure 3.3 – Les différents processus observés après l’attachement du VIH par DC-SIGN. [102]

### 3.1.1 Processus de présentation du virus par le passage dans des phagosomes

Le premier mécanisme est l'internalisation et la dégradation du virus dans des phagosomes qui conduit à une présentation « normale » d'antigène par le CMH. Mais durant cette étape, une partie des virus ne serait pas dégradée et présentée « intact » à la surface cellulaire facilitant la formation de la synapse virologique. Le virus détourne donc la capacité de reconnaissance d'une lectine de type C, DC-SIGN, pour être internalisé dans la cellule. Normalement l'internalisation par les PRRs conduit à la dégradation de l'agent pathogène dans des compartiments intracellulaires afin de présenter des antigènes aux lymphocytes T. Or, dans le cas du VIH, celui-ci peut être dirigé dans des corps multivesiculaires qui ne permettent pas la dégradation du virus mais sa présentation intact à la surface cellulaire (Figure 3.4, voie (ii))[26]. Cette présentation, via DC-SIGN, facilite la formation de la synapse infectieuse entre la cellule dendritique et le lymphocyte T [1]. Le virus peut donc infecter les lymphocytes T via les récepteurs CD4 et les co-récepteurs CCR5 ou CXCR4 (Figure 3.2).

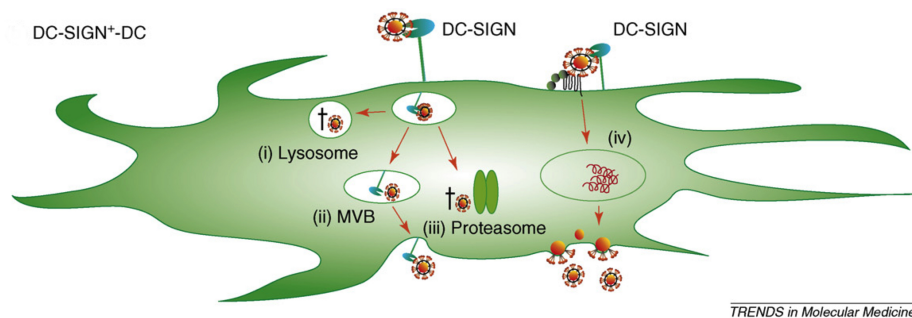


Figure 3.4 – Interaction DCs dermiques avec le VIH. Interaction du VIH avec les DCs dermiques via DC-SIGN : i) internalisation et dégradation du virus dans les lysosomes. ii) internalisation du virus dans des corps multivesiculaires conduisant à la présentation du virus intact à la surface cellulaire. iii) dégradation du virus par le protéasome. iv) infection directe des DCs par le virus (cis-infection).[26]

### 3.1.2 La concentration du virus à la surface des DCs sans internalisation

Un deuxième mécanisme impliquant DC-SIGN est le transport à la surface cellulaire de virus. En effet, il a été montré que le virus pouvait ne pas être internalisé mais se concentrer à la surface cellulaire par la diffusion latérale de DC-SIGN dans la membrane (phénomène appelé « surface trapping ») (Figure 3.2, extracellular transport)[12, 51]. Cette concentration permet de former des synapses virologiques entre la DC et le lymphocyte T facilitant l'infection de ces dernières.



### 3.1.3 La cis-infection

La trans-infection des lymphocytes T peut également être une conséquence d'une infection primaire des cellules dendritiques, qui s'appelle la cis-infection. Le virus peut ainsi infecter directement les cellules dendritiques par la reconnaissance de la glycoprotéine d'enveloppe du virus (gp120) par les récepteurs cellulaires CD4 et les co-récepteurs CCR5 (cis-infection, Figure 3.4, voie (iv)). En effet, les récepteurs CD4 et CCR5 sont le moyen utilisé par le virus pour infecter les lymphocytes T. Or, les cellules dendritiques possèdent également ce type de récepteur et elles peuvent donc être directement infectées par le virus. Cette infection conduit à la multiplication du virus dans les DCs durant leur migration vers les organes lymphoïdes secondaires. Les virions nouvellement produits peuvent donc directement infecter les lymphocytes T (trans-infection, Figure 3.4, voie (iv)). Cette cis-infection est facilitée par la reconnaissance du virus par DC-SIGN, qui fixe, concentre et permet sa transmission aux récepteurs CD4 et CCR5 (Figure 3.4). De plus, un article récent suggère que la fixation préalable à DC-SIGN augmenterait l'interaction avec CD4 [50].

## 3.2 Immuno-régulation par la reconnaissance du virus par DC-SIGN

En 2007, un article suggère que la reconnaissance du virus par DC-SIGN conduirait aussi à une voie de signalisation immunosuppressive. En effet, l'activation de DC-SIGN activerait la phosphorylation de NF $\kappa$ B (pour « nuclear factor kappa-light-chain-enhancer of activated B cells »), qui augmente et prolonge la transcription du gène de l'IL-10 (pour « Interleukine-10 »). L'IL-10 sécrétée inhibe la maturation et la présentation d'antigène par les DCs mais n'empêche pas la formation de la synapse immunologique avec les lymphocytes T [52, 90].

## 3.3 Accessibilité de DC-SIGN à la surface cellulaire comparée à CD4 et CCR5

L'accessibilité des récepteurs à la surface cellulaire joue un rôle important dans ces phénomènes de trans et cis-infection (Figure 3.5 A). En effet, DC-SIGN possède un domaine extracellulaire étendu lui facilitant l'accès aux virus et ainsi lui permettre de reconnaître avant les récepteurs CD4 et CCR5 impliqués dans la cis-infection. De plus, DC-SIGN, après son expression, est transportée à la membrane de façon à concentrer la protéine

### 3.4. UNE DÉFENSE NATURELLE CONTRE LE VIH : LES CELLULES DE LANGERHANS

dans certaines zones (en raft), lui augmentant le pouvoir fixateur de l'ensemble des molécules sur le virus (Figure 3.5 B). En effet, la disposition de DC-SIGN à la surface cellulaire en cluster augmente considérablement la capacité de fixation par un phénomène d'avidité de surface.

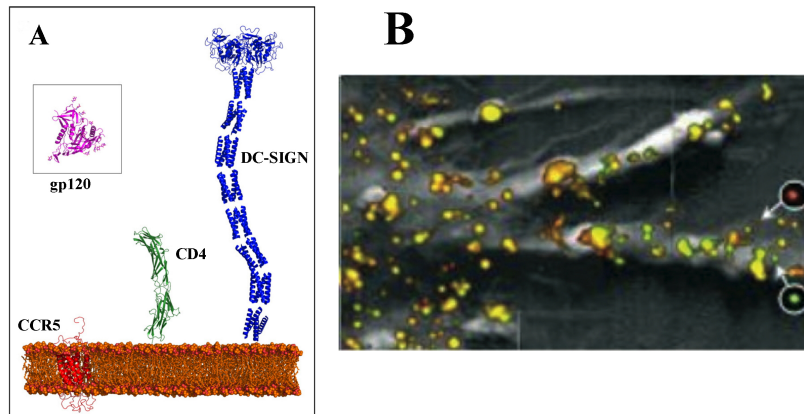


Figure 3.5 – A : Schéma représentant les tailles respectives des différents récepteurs, DC-SIGN, CD4 et CCR5. La structure utilisée pour DC-SIGN résulte d'une étude par SAXS [97], celle de CCR5 est un modèle (1ND8). B : Combinaison d'une topographie (gris) et d'images de fluorescences (couleur) [19]. La fluorescence représente les molécules de DC-SIGN à la surface cellulaire d'une cellule dendritique.

La trans-infection, sans passer par la cis-infection, est un mécanisme à court terme puisqu'il ne nécessite pas la production de nouveaux virions contrairement à la cis-infection qui demande quelques jours afin de produire de nouveaux virions. DC-SIGN joue un rôle majeur dans la transmission du virus aux lymphocytes T. Cette lectine de type-C, qui permet normalement d'induire une réponse immunitaire efficace, est ainsi détournée par le virus.

### 3.4 Une défense naturelle contre le VIH : les cellules de Langerhans

Tout d'abord, les cellules de Langerhans n'expriment pas DC-SIGN mais possèdent une autre lectine de type-C, la Langérine. Il a été démontré, en 2002 par Turville et al, que les cellules de Langerhans sont capables de reconnaître le VIH par le biais de la Langérine [103]. Environ 5 ans plus tard, Lot de Witte et al ont observé à la surface de cellules de Langerhans la co-localisation de la Langérine et du virus ainsi que l'internalisation de celui-ci dans les granules de Birbeck (Figure 3.6)[24].

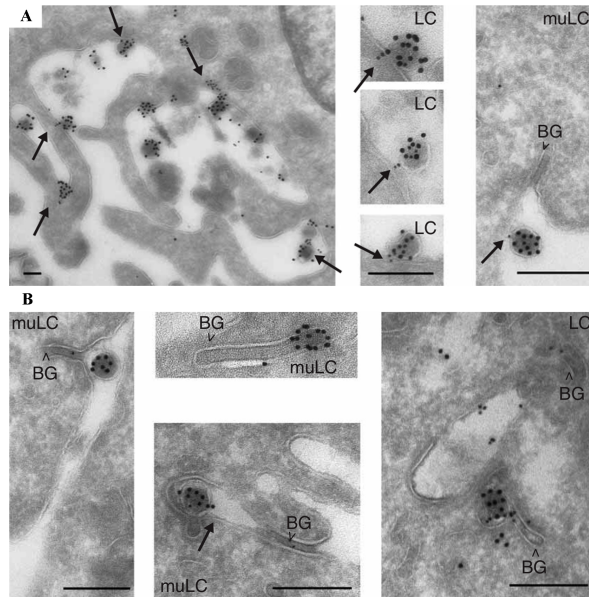


Figure 3.6 – Clichés de microscopie électronique de coupe de cellules de Langerhans en présence de VIH. La Langerine et la protéine p24 de capside du VIH sont marquées par des anticorps couplés avec des billes d’or de 10 nm et 15 nm respectivement. A) Co-localisation de la Langerine et du virus à la surface des cellules de Langerhans. B) Internalisation du virus dans les granules de Birbeck. La barre d’échelle représente 100 nm. [24]

La reconnaissance est réalisée par la faculté de la Langerine à reconnaître des motifs saccharidiques portés par la glycoprotéine d’enveloppe gp120 à la surface du virus. La reconnaissance du virus ne conduit pas à son internalisation dans une voie phagocytaire classique (contrairement à DC-SIGN) mais dans le granule de Birbeck. Cette différence d’acheminement du virus dans ces granules particuliers ne permet pas au virus d’être présenté intact à la surface cellulaire. Le devenir du virus reste encore inconnu mais une observation laisse penser que cette internalisation dans le granule de Birbeck est une « protection » vis-à-vis de l’infection par le VIH. En effet, lorsque le virus est localisé dans les BGs, l’infection des lymphocytes T n’est pas observée car la cellule de Langerhans n’est pas capable de présenter le virus à la surface cellulaire contrairement à DC-SIGN [24]. Il est à noter également que la cellule n’est elle-même pas infectée.

Mais cette protection peut être abolie. En effet, les LCs peuvent être converties en cellules favorisant la transmission du virus au lymphocyte T et cette conversion est strictement dépendante de l’expression de la Langerine ou de son inhibition [22, 61]. L’inhibition de la Langerine permet l’infection par les récepteurs CD4 et CCR5 des cellules de Langerhans et conduit donc à la production de nouveaux virions durant la migration des LCs vers les organes lymphoïdes secondaires et donc à la transmission de virus au lymphocyte T (Figure 3.7).

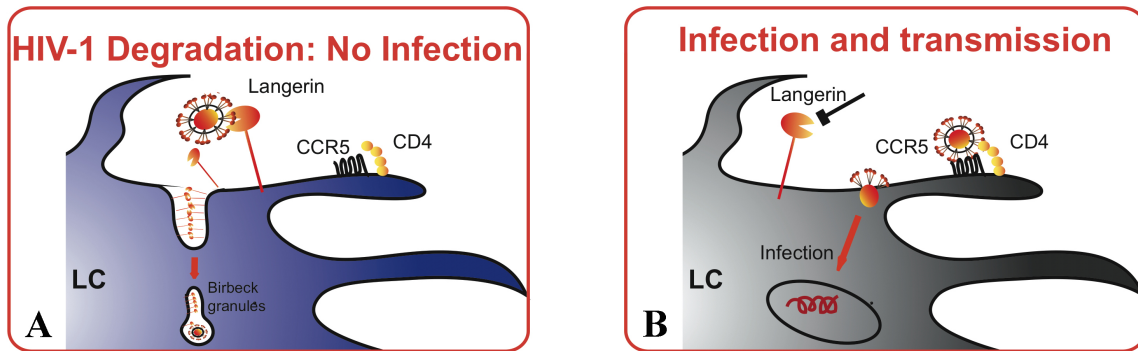


Figure 3.7 – Représentation schématique de l’interaction du virus avec de LCs en condition ”normale” (A) et lorsque la Langérine est inhibée. [24]

L’inhibition de la Langérine est possible par différents mécanismes :

- Saturation de la Langérine et infection directe des cellules de Langerhans via le récepteur CD4 et le co-récepteur CCR5 [61]. Lorsque la concentration virale est trop élevée, la Langérine est totalement saturée ce qui permet au virus en excès d’infecter directement les LCs par l’interaction avec le récepteur CD4 et le co-récepteur CCR5 [61].
- Activation des TLRs qui conduit à l’augmentation de la susceptibilité des cellules de Langerhans à être infectées[20, 22, 76]. En effet, la co-infection avec des virus (par exemple, l’herpès simplex 2)[22] ou par des bactéries Gram+ [76] conduit à l’activation des cellules de Langerhans et donc à la diminution de l’expression des PRRs, comme la Langérine. La perte de la Langérine, qui permet la protection des cellules de Langerhans vis à vis du VIH, à la surface de la cellule abolit l’effet protecteur de cette lectine.
- Toute forme d’inflammation des muqueuses produit une sécrétion de  $\text{TNF-}\alpha$  induisant à la diminution de l’expression de la Langérine et augmentant par conséquent la susceptibilité d’être infectée [20].

Dans ce phénomène d’élimination ou de transmission du virus, deux lectines de type-C jouent un rôle central. DC-SIGN permet la transmission du virus aux lymphocytes T alors que la Langérine induit l’internalisation de celui-ci dans les granules de Birbeck (Figure 3.8).

Ces deux lectines, très proches structuellement et fonctionnellement, sont néanmoins très différentes dans la signalisation cellulaire lors de la reconnaissance du virus. Une étude structurale et fonctionnelle de ces deux lectines est au coeur de la thématique abordée dans l’équipe afin de développer une stratégie de création de ligand spécifique de DC-SIGN sans pour autant inhiber la Langérine.

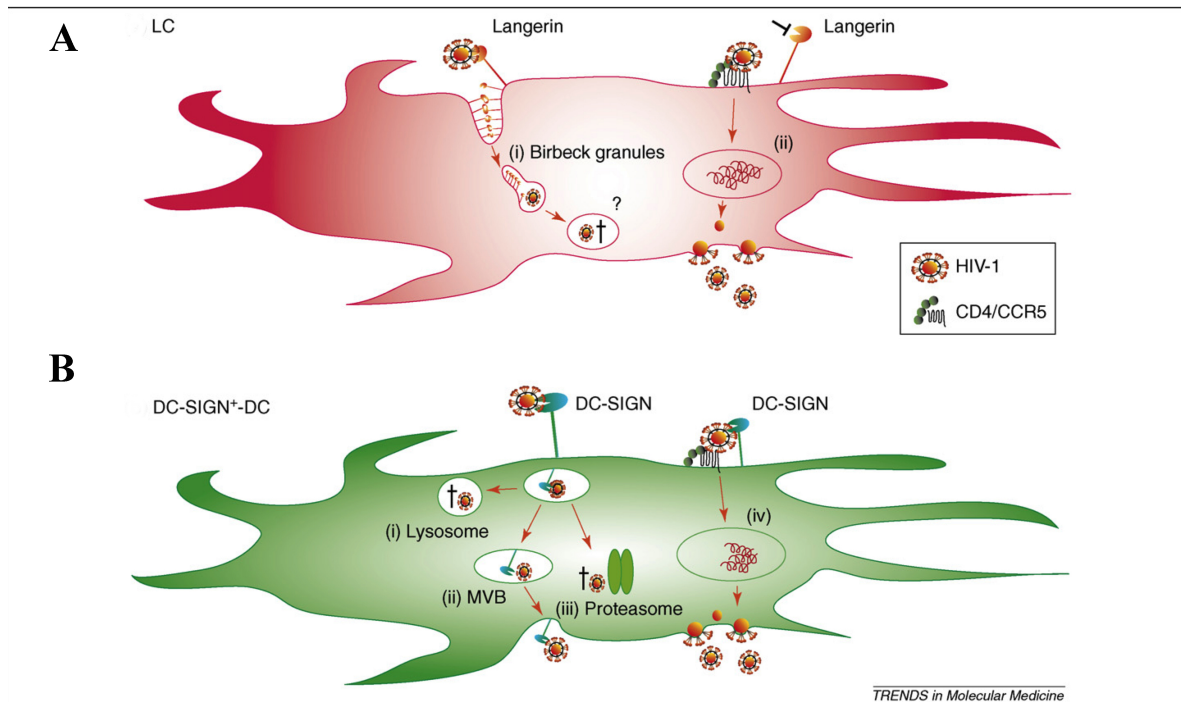


Figure 3.8 – Interaction DCs dermiques avec le VIH. a) Interaction du VIH avec les cellules de Langerhans : i) internalisation du virus dans les granules de birbeck qui permet de protéger l'infection. ii) L'inhibition de la Langérine (ou la saturation de la Langérine) conduit à l'infection des cellules de Langerhans qui permettra par le virion nouvellement formé d'infecter les lymphocytes T. b) Interaction du VIH avec les DCs dermiques via DC-SIGN : i) internalisation et dégradation du virus dans les lysosomes. ii) internalisation du virus dans des corps multivésiculaires conduisant à la présentation du virus intact à la surface cellulaire (Trans-infection). iii) dégradation du virus par le protéasome. iv) infection directe des DCs par le virus (Cis-infection).[26]

---

# Chapitre 4

## Les lectines

### 4.1 Informations générales sur les lectines

Les Lectines sont des protéines capables de reconnaître des oligosaccharides ou polysaccharides. Ces protéines sont retrouvées aussi bien chez les bactéries, les plantes et les animaux. La caractéristique première de ces protéines est la présence d'un domaine de reconnaissance des hydrates de carbone nommé CRD (« Carbohydrate recognition Domain »).

Les rôles physiologiques des polysaccharides sont nombreux. Ils peuvent avoir un rôle structural qui permet des interactions cellule-cellule mais aussi des interactions cellule-matrice extracellulaire. Au delà de site d'ancrage, ces hydrates de carbones complexes jouent des rôles importants dans des aspects de signalisation. Ceux-ci nécessitent la participation de lectines [111].

#### 4.1.1 Classification

Chez les animaux, on retrouve énormément de familles de lectines, plus précisément 13 familles de lectines. On peut regrouper ces différentes familles en deux principaux groupes :

- Les lectines intracellulaires
- Les lectines membranaires et extracellulaires

Les lectines intracellulaires sont composées de quatre familles principales : les calnexines et les lectines de types M, L et P. Ces lectines jouent des rôles essentiels dans le trafic intracellulaire, l'adressage des glycoprotéines ou encore dans la dégradation de celles-ci. Par exemple, les calnexines sont des protéines chaperonnes localisées dans le reticulum endoplasmique [18]. Elles permettent de vérifier la présence des bons oligosaccharides présents sur certaines protéines. Si celles-ci ne sont pas correctement glycosylées alors

elles ne sont pas exportées à l'appareil de Golgi pour finir leur maturation [18].

Les lectines extracellulaires concernent toutes les autres familles. On peut noter dans ce groupe de lectines, les lectines de type C et R, ainsi que les siglecs ou encore les galectines. Ces lectines sont généralement impliquées dans l'adhésion cellulaire, la signalisation cellulaire, la clairance de glycoprotéine ou encore dans la reconnaissance de pathogènes (Tableau 4.1). Par exemple, CD22, une lectine de la famille des siglecs, est impliquée dans la régulation négative de l'activité des lymphocytes B [48].

Par la suite, nous développerons les lectines de type C. Cette famille est très abondante chez les mammifères et possède des organisations très variées. Ces organisations varient dans le nombre de CRD exposé, et/ou la façon dont ceux-ci s'oligomérisent pour augmenter l'affinité par exemple.

### 4.1.2 Reconnaissance des ligands

Une caractéristique prédominante des lectines est la faible affinité des CRDs pour leurs ligands. En effet, chez les mammifères les lectines ont une affinité pour des monosaccharides de l'ordre du mM. Cette faible affinité ne permet pas de fixer fortement leurs ligands pour induire un signal. Pour effectuer correctement leurs rôles, les lectines utilisent différents mécanismes permettant d'augmenter l'affinité par des phénomènes de multivalence. Cette multivalence peut être induite par l'oligomérisation de la protéine, par la présence de plusieurs CRDs sur celle-ci, ou encore par de la multivalence de surface pour les lectines membranaires ne possédant qu'un seul CRD et incapables de s'oligomériser (Figure 4.1). On peut retrouver aussi une combinaison de plusieurs facteurs comme l'oligomérisation de la protéine ainsi que de la multivalence de surface.

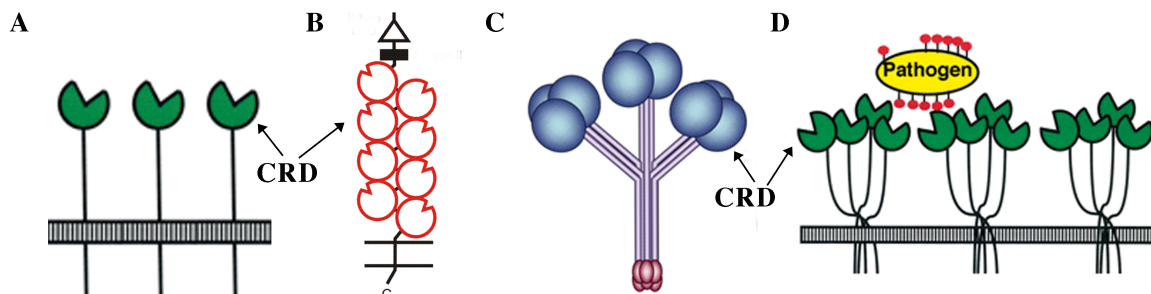


Figure 4.1 – Représentation schématique des plusieurs modes d'avidités engendrés par les lectines. A : multivalence de surface. B : multivalence par la présence de plusieurs CRDs au sein d'une même chaîne. C : multivalence par oligomérisation de la protéine. D : association de la multivalence de surface et de l'oligomérisation.



Famille de lectine	Ligands typiques	Localisation cellulaire	Exemples de fonctions
Calnexin	Glc1Man9	RE	Gestion des protéines dans le RE
Lectines de type-M	Man8	RE	Dégradation de glycoprotéines
Lectines de type-L	Variés	RE, Golgi	Gestion des protéines dans le RE
Lectines de type-P	Man 6-phosphate, autres	Voie de sécrétion	Gestion des protéines post-Golgi, adressage des glycoprotéines
Lectines de type-C	Variés	Membrane cellulaire, extracellulaire	Adhésion cellulaire (selectins), immunité innée (collectins)
Galectines	b-Galactosides	Cytoplasme, extracellulaire	Agent de pontage de glycannes dans la matrice extracellulaire
Lectines de type-I (siglecs)	Acide Sialique	Membrane cellulaire	
Lectines de type-R	Variés	Golgi, membrane cellulaire	Ciblage d'enzymes
Lectines F-box	GlcNAc2	Cytoplasme	Dégradation de glycoprotéines non structurées
Ficolines	GlcNAc, GalNAc	Membrane cellulaire, extracellulaire	Immunité innée
Lectines de type Chitinase	Chito-oligosaccharides	Extracellulaire	Métabolisme du collagène
Lectines de type-F	Oligosaccharides terminant par un Fuc	Extracellulaire	Immunité innée
Intelectines	Gal, galactofuranose, pentoses	Extracellulaire/membrane cellulaire	Immunité innée, fertilisation et embryogénèse

Tableau 4.1 – Liste des différentes familles de lectines, de leurs ligands, leurs localisations cellulaires et de leurs fonctions. (Drickamer, tableau traduit de <http://www.imperial.ac.uk/research/animallectins/ctld/lectins.html>)

## 4.2 La famille des lectines de type C

### 4.2.1 Organisation structurale

La caractéristique commune, de toutes les lectines de cette famille est un domaine de type CRD (« Carbohydrate Recognition Domain »). Celui-ci possède un ion calcium participant à la reconnaissance des sucres. Cet ion intervient aussi dans l'organisation structurale de ce domaine. Il est donc essentiel pour le fonctionnement de ces protéines. La particularité des CRDs est que leurs structures sont très conservées et particulièrement du point de vue de certains ponts disulfures essentiels dans le maintien de cette structure (Figure 4.2). En effet, on retrouve au minimum chez ces lectines au moins deux ponts disulfures avec le pont disulfure 1 qui est présent chez toutes ces protéines. Ces lectines sont en général composées de deux hélices alpha ( $\alpha 1$  et  $\alpha 2$ ) et de cinq brins  $\beta$  ( $\beta 1-5$ ) ainsi que de quatre boucles (L1 à L4) [31]. Ces lectines peuvent aussi former des homo ou hétérodimères par des ponts disulfures dus à la présence de cystéine aux positions 7, 8 et 9.

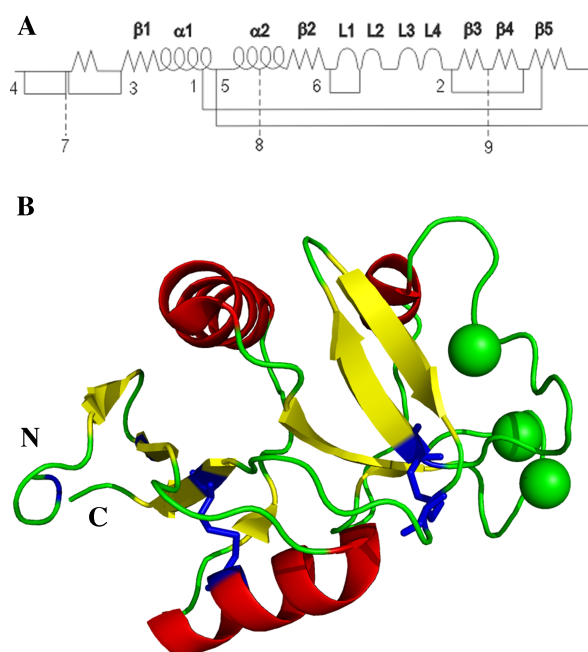
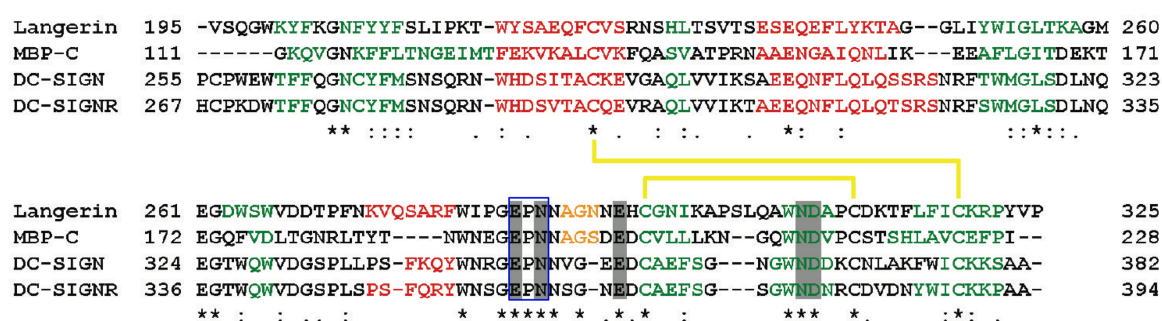


Figure 4.2 – A : représentation de la structure secondaire des CTLDs. Les ponts disulfures possibles dans ce type de lectine sont représentés par les lignes numérotées. B : Structure cristallographique de DC-SIGN. En bleus sont représentés les ponts disulfures (PDB id : 2XR6). En verts les ions calciums. <http://www.imperial.ac.uk/research/animallectins/ctlld/disulfides.html>

Les CRDs sont tous les domaines ressemblant aux lectines de type-C, mais dans le contexte d'étude de lectine de type-C où ils ont une activité avérée de reconnaissance des

sucres, nous utiliserons le terme CRD pour « carbohydrate recognition domain ».

Les structures des CRDs de ces lectines sont très similaires voir identiques (Figure 4.3). On peut trouver chez ces lectines plusieurs sites calcium, mais le site 1 est toujours conservé chez ces protéines.



## 4.2.2 Spécificité pour la reconnaissance des saccharides

Les lectines de type-C possèdent deux motifs principaux permettant la coordination du calcium. Ces deux motifs, EPN et QPD, imposent une organisation particulière de la boucle en coordination avec l'ion calcium et ainsi imposent une spécificité de reconnaissance. En effet, le motif EPN procure une spécificité de reconnaissance pour le mannose, N-acetylglucosamine et le fucose alors que le motif QPD procure lui une spécificité de reconnaissance pour le galactose [64](Figure 4.4).

Ce type de lectine possédant un motif EPN nécessite une orientation précise des groupes hydroxyles des sucres. En effet, ce motif impose que les groupes hydroxyles en position 3 et 4 du sucre soient orientés en équatorial (Figure 4.4 A). Effectivement, toutes les structures, actuellement disponibles dans la banque de données « Protein Data Bank », du CRD de DC-SIGN et de la Langérine en présence d'oligosaccharide montrent une seule configuration du sucre en liaison avec le calcium (code PDB : 3P5H, 3P5F, 3P5E, 3P5D, 3P5G, 3P7G, 3P7H, 1K9I, 1SL4, 2IT5, 2IT6, 2XR5).

Contrairement aux lectines de type-C possédant le motif EPN, celles liant l'ion calcium par le motif QPD nécessitent une orientation différente du groupement hydroxyle en position 4 du saccharide. En effet, il est nécessaire que le saccharide possède ce groupement hydroxyle en axial pour effectuer une interaction (Figure 4.4 B). Cette orientation est retrouvée dans le galactose. Chez les lectines de type-C à motif QPD, d'autres résidus sont

conservés et en particulier un tryptophane qui permet une interaction avec les carbones du cycle osidique. Donc, le mode de coordination du calcium couplé à la présence du tryptophane impose une orientation particulière du sucre qui ne peut être possible que si les hydroxyles C3 et C4 sont en position équatoriale et axiale respectivement [64].

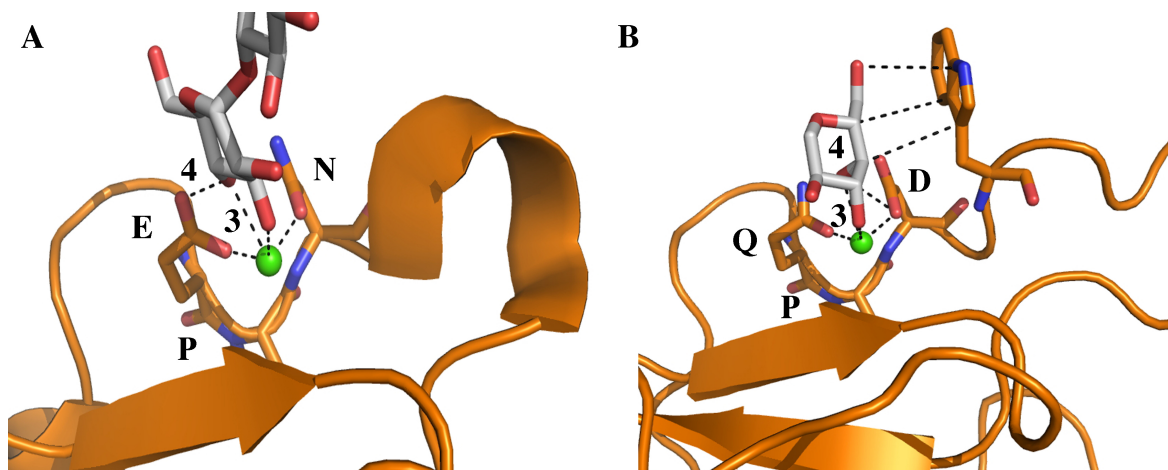


Figure 4.4 – Structure cristallographique de lectine de type-C possédant le motif EPN (A) et QPD (B). A : structure cristallographique de la Langérine en présence de maltose (3P7H) [14]. B : structure cristallographique du récepteur scavenger en présence de galactose (2OX9) [37].

### 4.2.3 Classification

Les lectines de type C se répartissent entre des protéines associées à la membrane cellulaire et des protéines sécrétées sous forme solubles. Elles participent notamment à l'adhésion cellulaire ou encore à la reconnaissance de pathogènes (ou les deux) ainsi qu'aux fonctions de migration cellulaire. Actuellement, cette famille de lectine est sous-divisée en 14 groupes (Figure 4.5).

Parmi celles qui sont transmembranaires, on peut identifier deux types de lectines. Le type 1 inclut les lectines transmembranaires exposant le côté C-terminal de leur séquence protéique dans le cytoplasme, et par conséquent le type 2 est représenté par les protéines orientant le côté N-terminal dans le cytoplasme. Dans ces type de lectines transmembranaires, on retrouve les selectines (type 1) ou encore DC-SIGN et la Langérine (type 2).

On retrouve 7 groupes dans les lectines transmembranaires de type 1. Parmi eux, il y a notamment le groupe VIII (ou récepteurs de type 1) avec les chondrolectines exprimées par les cellules musculaires [112] ou encore le groupe XI avec les attractines responsables des interactions des lymphocytes T avec les macrophages ou les cellules dendritiques [32]. Tous ces groupes de lectines ne seront pas développés.

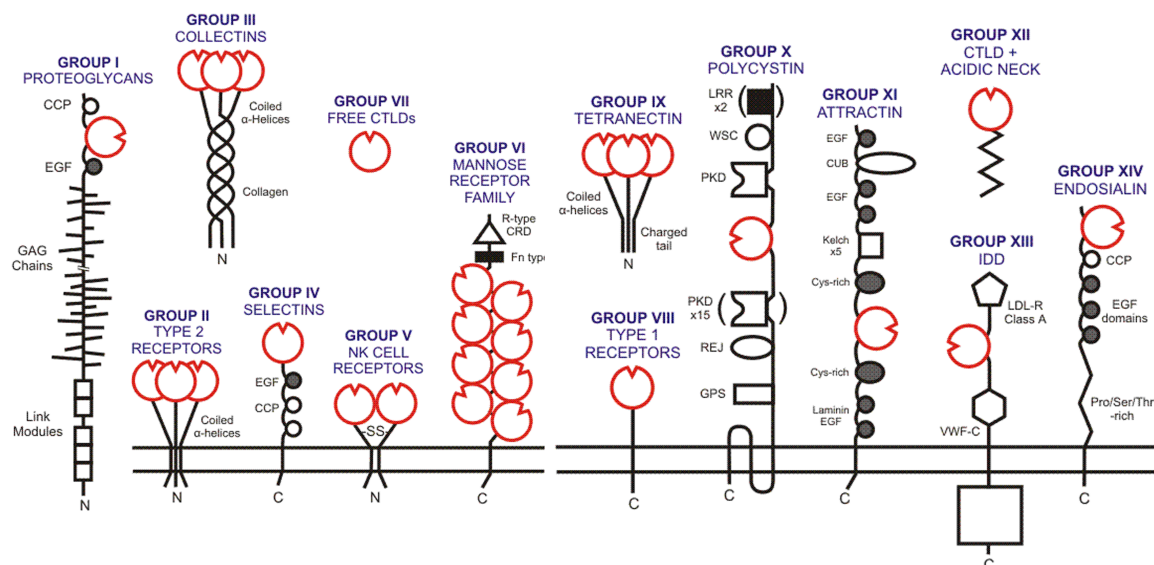


Figure 4.5 – Les différents groupes de lectines de type C. Image réalisée par le Prof. Drickamer <http://www.imperial.ac.uk/research/animallectins/ctld/lectins.html>

Les récepteurs lectines transmembranaires de type 2 sont composés de 2 groupes. Le groupe II (ou récepteurs de type 2) est composé de récepteurs exprimés à la surface des macrophages et des cellules dendritiques entre autres. La particularité de ce type de lectines est l'oligomérisation de celles-ci par un domaine coiled-coil composé d'hélices alpha. Cette oligomérisation permet d'augmenter l'affinité globale pour un ligand par la présentation de plusieurs CTLD par une même protéine. Ces récepteurs sont impliqués généralement dans la reconnaissance de pathogène avec l'exemple de DC-SIGN qui est capable de reconnaître un ensemble de différents pathogènes comme le VIH [35].

Le deuxième groupe composant ces lectines est le groupe des récepteurs des cellules NK (« Natural Killer », groupe V). Ces récepteurs sont impliqués dans des mécanismes de l'immunité innée, comme par exemple le récepteur NKG2D qui joue un rôle majeur dans l'activation des cellules NK et le déclenchement de la production de cytokines pro-inflammatoire, dans la défense contre une infection par *Pseudomonas aeruginosa* par exemple [113].

Par la suite, nous développerons seulement le groupe II avec les lectines DC-SIGN et la Langérine.

### 4.3 Les lectines de type C du groupe II

Ces lectines sont des capteurs de glycosylation portés par des agents pathogènes [28, 82], mais elles peuvent avoir d'autres fonctions associées comme l'adhésion cellulaire [47] ou la reconnaissance de glycoprotéine des lymphocytes T [98] en passant par la formation

d'organite cellulaire [104].

### 4.3.1 Organisation structurale

Les lectines de ce groupe possèdent une organisation similaire (Figure 4.6). Elles sont composées :

- D'un domaine intracellulaire (côté N-terminal) permettant la transduction du signal.
- D'un domaine transmembranaire.
- D'un domaine coiled-coil composé d'hélices alpha permettant l'oligomérisation (Neck).
- D'un domaine de reconnaissance des hydrates de carbones (CRD).

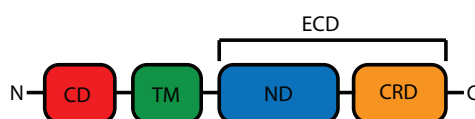


Figure 4.6 – Représentation schématique de l'organisation générale des lectines de type C du groupe II. CD : domaine cytoplasmique (cytoplasmic domain); TM : domaine transmembranaire (transmembrane domain); ND : domaine neck (neck domain); CRD : domaine de reconnaissance des sucres; ECD : domaine extracellulaire composé du domaine neck et du CRD.

Le domaine cytoplasmique peut contenir des séquences permettant l'internalisation du récepteur ou encore des séquences activant des voies de signalisation. Par exemple, l'interaction du VIH avec le récepteur DC-SIGN induit une voie de signalisation dépendante de RAF-1 qui finit par l'activation de la transcription de l'IL-10. La production de l'IL-10 conduit à une immunosuppression [102].

Le domaine transmembranaire est essentiellement composé de résidus hydrophobes. Ce domaine est court (environ 20 acides aminés) et permet de traverser la membrane plasmique.

Le domaine extracellulaire est très allongé et permet de présenter le CRD dans la matrice extracellulaire afin d'analyser l'environnement. Le domaine neck permet l'oligomérisation de la protéine pour former le complexe quaternaire. Ce domaine forme une structure de type coiled-coil formée d'hélices alpha. On retrouve des coiled-coil dimérique, trimérique ou encore tétramérique.

### 4.3.2 Le domaine coiled-coil

Leur mode d'organisation oligomérique de ces protéines conduit à de l'avidité par de la multivalence dans les interactions.

Un domaine coiled-coil est composé de l'association d'hélices alpha qui s'enroulent les unes sur les autres selon une stoechiométrie pouvant aller de 2 à 4 [69]. L'organisation en

coiled-coil nécessite un motif répété d'acides aminés organisé en heptade (pour les coiled-coil parallèles). Ces heptades d'acides aminés sont désignés par les lettres **abcdefg**. En tenant compte du pas de l'hélice, les résidus **a** et **d**, **principalement hydrophobes**, se trouvent sur la même face et ils sont essentiels dans la formation du coeur hydrophobe du coiled-coil (Figure 4.7)[120, 119].

Les positions **e** et **g** sont, quand à elles, composées de résidus chargés permettant de stabiliser ou déstabiliser la formation du coiled-coil mais aussi d'être responsables de l'organisation parallèle ou anti-parallèle. Dans les modèles présentés (Figure 4.8), le coeur hydrophobe est essentiel à l'organisation du coiled-coil.

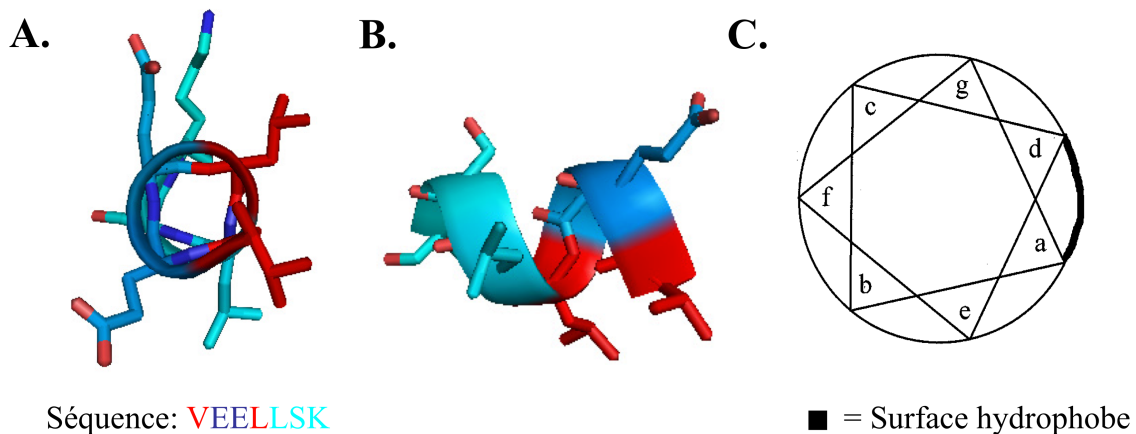


Figure 4.7 – Représentation d'un heptade d'acides aminés formant une hélice  $\alpha$  de type coiled-coil. A : vue du dessus. B : vue latérale. C : schéma de distribution des acides aminés d'une heptade.

On retrouve aussi bien des structures dimériques, trimériques ou tétramériques de ce type d'organisation. Comme illustré sur la Figure 4.8, les résidus **a,d,e,g** sont essentiels dans le maintien de cette structure.

La longueur du domaine neck de ces lectines dépend exclusivement du nombre d'heptades présent dans la séquence de ce domaine. En effet, DC-SIGN par exemple possède 8 répétitions de ce motif alors que la Langérine n'en possède que 5. La longueur du neck est importante pour ces lectines car elles sont exposées à la surface cellulaire et elles sont principalement des PRRs. Ces domaines neck permettent de projeter loin de la surface cellulaire les CRDs augmentant ainsi leur efficacité. Donc pour effectuer correctement leur rôle, plus elle dépasse de la surface cellulaire plus elles ont de chance d'accéder à un antigène donné et de le reconnaître. Par exemple dans le cas du VIH présenté précédemment, l'infection des cellules dendritiques par les récepteurs CCR5 et CD4 est difficile par la différence de taille avec DC-SIGN (Figure 3.4 A).



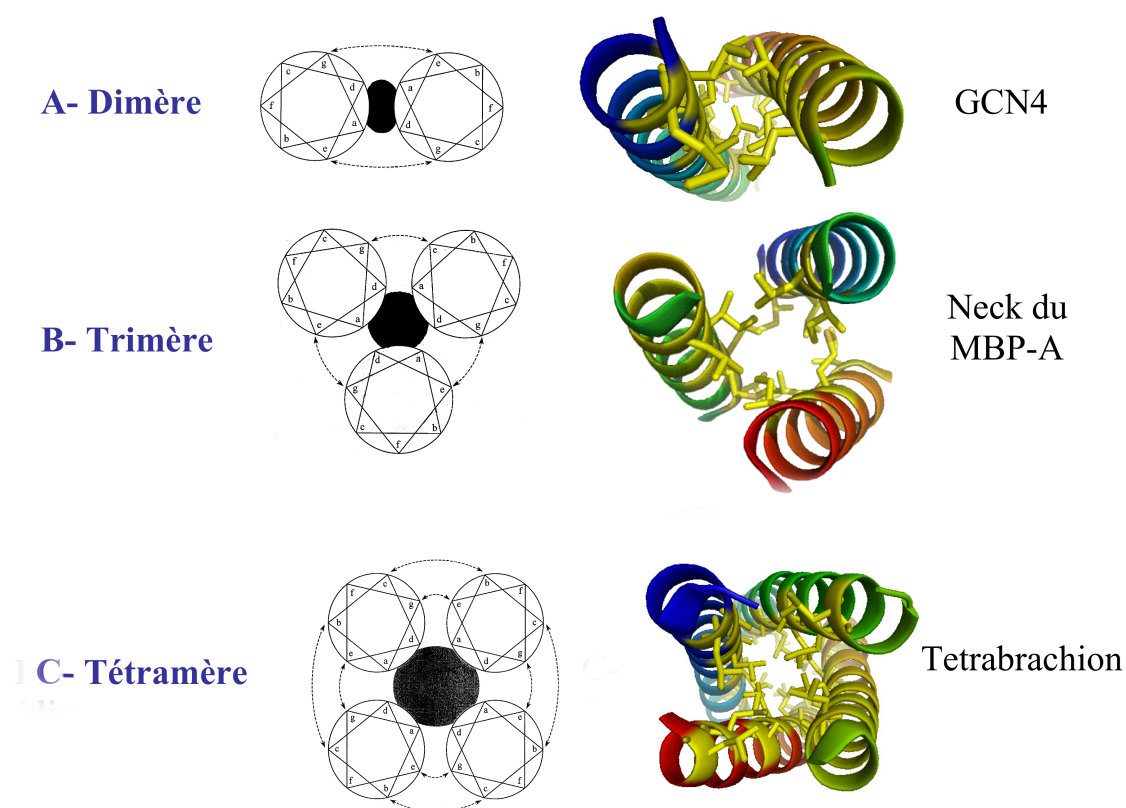


Figure 4.8 – Quatre modèles d'organisation oligomérique de type coiled-coil parallèle avec des exemples de structures résolues. A : Un dimère de GCN4 (PDB : 2ZTA, [78]). B : Un trimère de MBP-A (PDB : 1RTM, [110]). C : Un tétramère de Tetrabrachion (PDB : 1FE6, [96]).



# Chapitre 5

## La Langérine

### 5.1 Structure de Langérine

#### 5.1.1 Vue d'ensemble de la Langérine

La Langérine est composée de trois principaux domaines caractéristiques des lectines du groupe II (Figure 5.1)[104]. Le domaine cytoplasmique (CD), du côté N-terminal, possède un domaine riche en proline permettant l'export de la protéine à la membrane [100]. Un domaine transmembranaire (TMD) court permet de présenter à la surface un long domaine extracellulaire (ECD). Ce dernier est composé d'un domaine riche en leucine et permet l'oligomérisation de la protéine par un coiled-coil d'hélices  $\alpha$ . Enfin le domaine de reconnaissance des hydrates de carbones porte le motif EPN (Figure 5.1). L'étude de la Langérine recombinante a permis de démontrer un état stable trimérique du domaine extracellulaire[92, 100]. L'oligomérisation de ces lectines est très importante pour l'interaction avec des ligands.

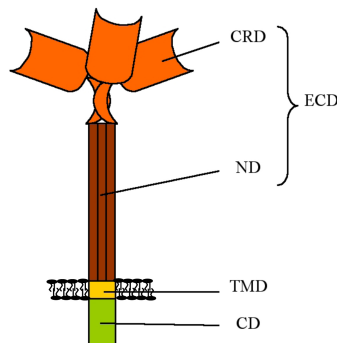


Figure 5.1 – Schéma représentant l'architecture de la Langérine. CD : domaine cytoplasmique. TMD : domaine transmembranaire. ND : domaine neck. CTLD : domaine de lectine de type-C. ECD : domaine extracellulaire. [106]

Comme présenté précédemment, le motif EPN procure une spécificité de reconnaissance pour les saccharides possédant des mannoses, N-acetyl-glucosamine et le fucose.

### 5.1.2 Implication de la Langérine dans le granule de Birbeck

Comme évoqué précédemment, les études menées par Valladeau et al ont montré le rôle indispensable de la Langérine dans la formation et la structuration des granules de Birbeck. Ainsi, une étude complémentaire réalisée par l'équipe en collaboration avec Jenny Valladeau et Colette Dezutter-Dambiant a été menée pour approfondir les connaissances sur l'implication de la Langérine dans le granule de Birbeck. Nous avons ainsi montré que le domaine de reconnaissance des hydrates de carbone (CRD) de la lectine est nécessaire à la formation des BGs (Figure 5.2, A). Par sa déplétion, la formation de BGs est abolie montrant le rôle central du CRD de la Langérine dans l'établissement de structure à membrane zippée. Le domaine cytoplasmique de cette lectine joue un rôle important dans l'adressage de la protéine à la membrane (Figure 5.2, B). La suppression ou la modification de ce domaine conduit à l'accumulation de Langérine dans une structure évoquant le reticulum endoplasmique (par la forte présence de ribosome) ou à la formation de sandwich cytomembranaire (CMS) connecté à des corps multivésiculaires (Figure 5.2, C).

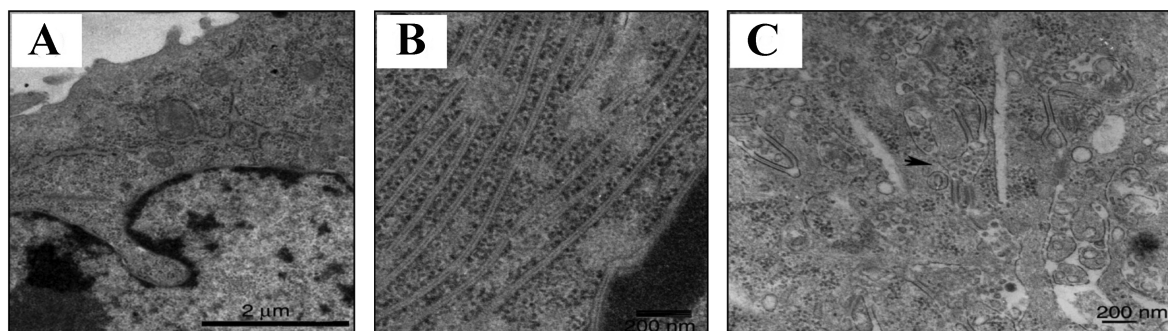


Figure 5.2 – Micrographies des différentes constructions de la Langérine. A : transfection réalisée avec le gène de la Langérine déplété du CRD, B : transfection réalisée avec le gène de la Langérine déplété du domaine cytoplasmique, C : transfection réalisée avec le gène de la Langérine contenant des mutations dans le domaine cytoplasmique. [100]

Une analyse comparative de microscopie électronique des granules de Birbeck couplée à la modélisation de la partie extracellulaire de la Langérine, nous ont permis de proposer deux modèles d'organisation de la protéine dans le granule de Birbeck ainsi qu'un modèle d'invagination des membranes (Figure 5.3)[100]. En effet, à partir de l'estimation de la distance entre les deux membranes (25-40 nm) et de l'étude de l'organisation de la partie extracellulaire de la Langérine, il a été proposé que la Langérine s'organise soit « tête-

bêche » soit en alternance, pouvant expliquer la densité centrale observée par microscopie électronique (Figure 5.3B-C). Il a été suggéré ainsi un modèle d'invagination de membrane impliquant directement la Langérine (Figure 5.3A-D).

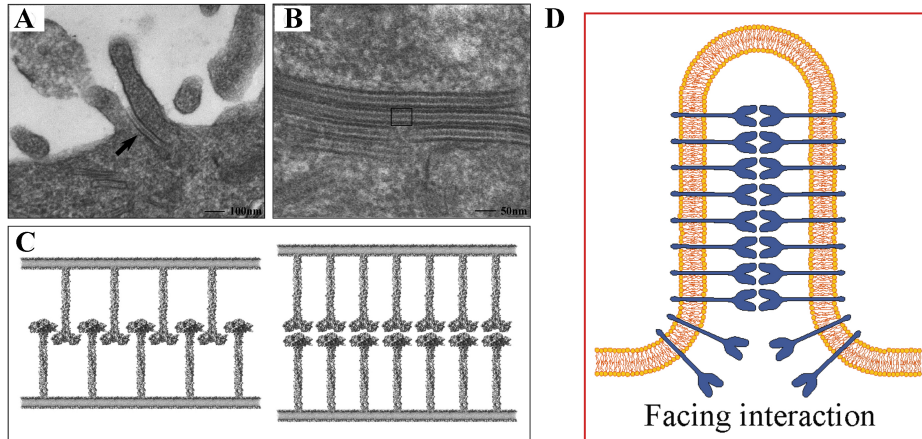


Figure 5.3 – Les granules de Birbeck dans les cellules de Langerhans et les modèles de l'organisation macromoléculaire. A : Invagination de la membrane plasmique formant une structure en sandwich de membrane dans une cellule de Langerhans. B : Granules de Birbeck des cellules de Langerhans. C : Deux modèles de l'organisation macromoléculaire de la Langérine dans le granule de Biebeck. D : Modèle de la déformation de la membrane plasmique par la Langérine. [100]

Actuellement, une seule autre protéine a été co-localisée dans les granules de Birbeck, CD1a. Ce partenaire, CD1a, est une glycoprotéine caractérisée par sa faculté à présenter des antigènes non-peptidique (lipidique, et glycolipidique) aux Lymphocytes T [68, 33]. Malgré le fait que le manque de granule de Birbeck n'affecte pas la fonction de présentation d'antigènes des cellules de Langerhans, cette co-expression de la Langérine et de CD1a suggère un possible rôle des granules de Birbeck dans des mécanismes alternatifs de présentation d'antigène par le granule de Birbeck [33].

### 5.1.3 Structure du CRD

Le CRD de la Langérine possède une structure similaire aux autres lectines de type-C. En effet, elle possède les 2 hélices  $\alpha$ , les 5 brins  $\beta$ , et les deux ponts disulfures minimums qui caractérisent les lectines de type-C (Figure 5.4 A). Contrairement à d'autres lectines de type-C de cellules dendritiques (DC-SIGN et la MBP), la Langérine possède un seul site calcique au sein de sa structure (site 1, Figure 5.4 B). On observe une orientation identique des chaînes latérales des résidus en coordination avec l'ion calcium, pour les trois structures, qui maintiennent l'intégrité du site 1 (Figure 5.4 D). Pour les autres sites, la non conservation de la séquence protéique de la Langérine implique dans ces régions la

modification de la position de structure secondaire dans les sites 2 et 3 (Figure 5.4 C) ainsi que de la perte de coordonnant pour le calcium dans le site 4 (Figure 5.4 E).

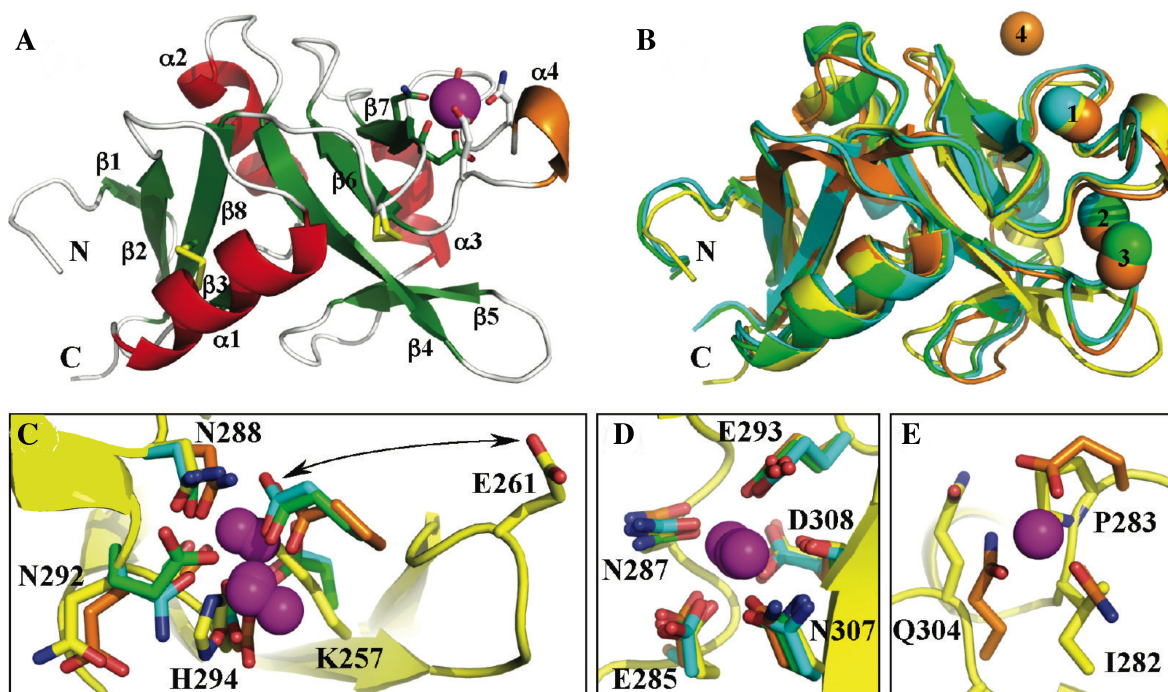


Figure 5.4 – Structure cristallographique du CRD de la Langérine (A) alignée avec celles de DC-SIGN (vert), de la MBP (orange) et de DC-SIGNR (cyan) (B). C : Zoom sur l’environnement du site calcium 2. D : Zoom du site calcium 1. E : Zoom du site 4.

### 5.1.4 Structure de l’ECD

La structure cristallographique d’une forme trimérique de la Langérine a été publiée pendant ce travail de thèse (PDB code : 3KQG). Celle-ci a pu être réalisée par la troncation des 2/3 du ND (Figure 5.5). Cette région de la protéine est prédominante pour l’oligomérisation de la protéine et l’orientation spécifique donnée aux CRDs.

On retrouve 4 interfaces différentes :

- La première qui constitue les interactions engendrées par le coiled-coil des hélices  $\alpha$  du ND (Figure 5.5 A). Cette interface est composée principalement d’interactions hydrophobes aux positions a et d de l’heptade (Figure 4.7).
- La seconde interface est composée du CRD d’un monomère et du ND d’un autre monomère (Figure 5.5 B). Plusieurs interactions sont observées entre la boucle, qui connecte le ND du CRD d’un monomère, et le brin  $\beta$  et l’hélice  $\alpha$  du monomère voisin. On retrouve des interactions de Van der Waals des résidus Gly198, Ser196 et Gln197 du ND avec respectivement les résidus Lys200, Tyr201 et Phe241 du CRD voisin. Des liaisons hydrogènes sont effectuées entre la Ser196 et la Tyr 201.

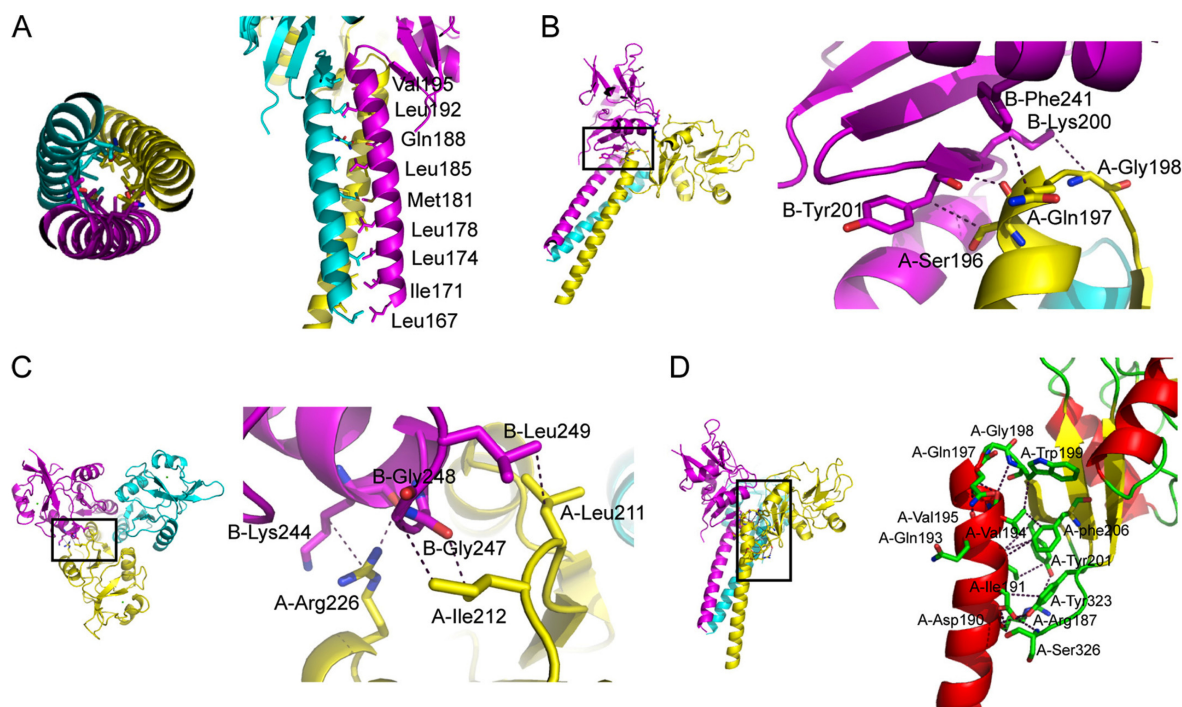


Figure 5.5 – Structure cristallographique d’une partie du domaine extracellulaire de la Langérine (PDB : 3KQG). A : coiled-coil vue du dessus et de côté. B : Interface ND d’un monomère avec le CRD du monomère voisin. C : interface CRD/CRD. D : interface ND et CRD d’un même monomère. [38]

- La troisième est constituée par des interactions entre les CRDs voisins (Figure 5.3 C). Ces interactions sont engendrées par la proximité des résidus Leu211 et Leu249, Ile212 et Gly247 ainsi que de Arg226 avec Gly248 et Lys244.
- La dernière interface fait intervenir les résidus de l’hélice  $\alpha$  du ND avec les résidus du CRD d’un même monomère (Figure 5.5 D). Cette interface implique un réseau d’interaction non polaire complexe.

La trimérisation permet une augmentation de la capacité de fixation de ce récepteur et conduit probablement à une spécificité de reconnaissance.

## 5.2 Reconnaissance et mode d’interaction de la Langérine

### 5.2.1 Introduction et nomenclature usuelle des sucres

Les hydrates de carbone sont l’appellation historique utilisée pour les glucides. Cette appellation provient des premières découvertes qui caractérisaient les glucides comme des structures de formule brute  $C_n(H_2O)_n$ . Or, tous les glucides ne sont pas composés que

d'atomes de carbone, d'oxygène et hydrogène, on retrouve aussi des structures glucidiques possédant des atomes d'azote, de phosphore ou de soufre. Comme évoqué précédemment, la Langérine reconnaît les glucides de type hexoses comme le glucose et le mannose. Différentes représentations des structures saccharidiques sont possibles mais nous n'utiliserons que la représentation de Haworth, qui permet de visualiser les structures saccharidiques cyclisées, ainsi qu'une nomenclature plus simple proposée par le « Consortium for Functional Glycomics » (CFG). La majorité des sucres considérés dans ce travail sont des aldohexoses existant essentiellement en solution sous la forme cyclisée par la réaction de l'hydroxyle porté par le carbone 5 avec la fonction aldéhyde du carbone 1 qui sera nommé carbone anomérique (Figure 5.6). Le carbone anomérique existe sous deux configurations différentes, soit la conformation alpha, soit la conformation bêta suivant la position du groupement hydroxyle par rapport à la position du carbone 6.

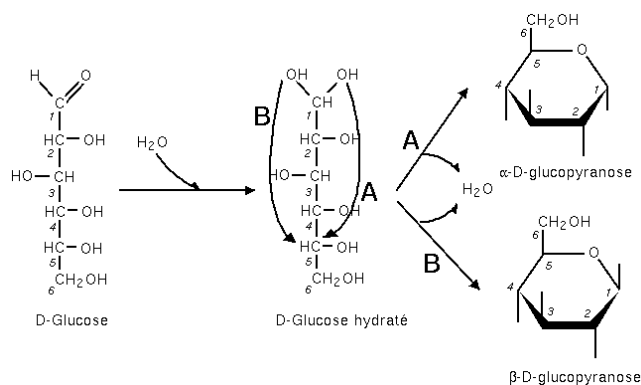


Figure 5.6 – Cyclisation du D-glucose.

Dans un souci de simplification pour des structures oligosaccharidiques ou polysaccharidiques, qui sont composées d'un enchaînement de plusieurs monosaccharides (de 1 à 10 pour les oligosaccharides et supérieur à 10 pour les polysaccharides), nous allons utiliser la nomenclature symbolique proposée par le CFG (Figure 5.7). Dans cette nomenclature, on trouve différentes formes comme le rond, le carré, le triangle et le losange. Le rond et le carré représentent respectivement l'hexose et l'hexosamine (ainsi que le N-acetyl-hexosamine), le triangle est utilisé pour le fucose et les losanges pour les saccharides possédant une fonction acide. Un code couleur définit la nature du sucre. Ainsi les formes dérivées du galactose sont représentées en jaune, le glucose en bleu, le mannose en vert et le fucose en rouge. Les carrés pleins symbolisent les N-acetyl-hexosamines, les partiellement pleins les hexosamines. Pour les saccharides acides, représentés sous la forme de losange, nous allons utiliser le losange partiellement colorié de bleu représentant l'acide glucuronique et son épimère en 5, l'acide iduronique est lui partiellement colorié en marron.



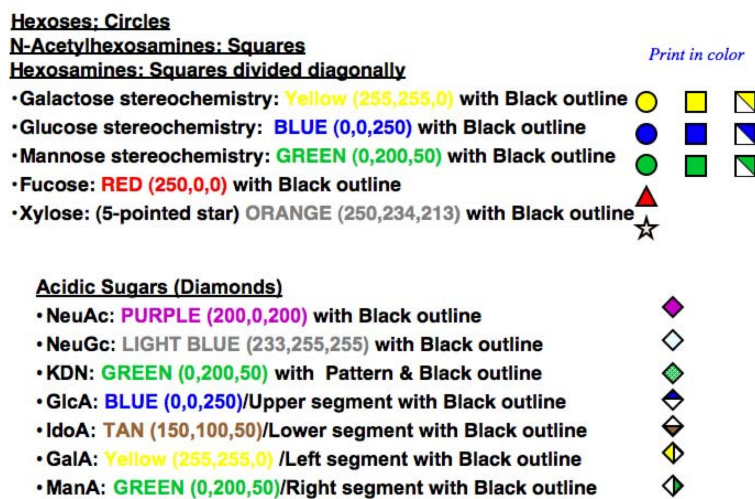


Figure 5.7 – Nomenclature proposée par le consortium de glycobiologie.

Dans la suite de ce chapitre, ne seront évoquées que les structures saccharidiques utilisées dans ce travail de thèse c'est-à-dire les structures de types « High Mannose » et de type glycosaminoglycanes.

### 5.2.2 Reconnaissance de la Langérine

L'affinité du CTLD de la Langérine pour des monosaccharides est seulement de l'ordre du millimolaire. Ainsi, le CTLD a une affinité de l'ordre de 2,5 mM pour le mannose, N-acetyl-glucosamine et le fucose (Tableau 5.1) [92].

Monosaccharide	$K_i$ (mM)	$K_{i,\text{Sugar}}/K_{i,\text{Man}}$
Mannose	$2.3 \pm 0.05$	1
N-acetylglucosamine	$2.8 \pm 0.01$	1.2
L-fucose	$2.6 \pm 0.5$	1.1
Galactose	$16.1 \pm 1.8$	7.0
$\alpha$ -Methylmannoside	$1.7 \pm 0.2$	0.7
$\alpha$ -Methylgalactoside	> 200	> 100
Glucose	$7.0 \pm 0.9$	3.0
N-acetylmannosamine	$5.6 \pm 0.9$	2.4
Man <sub>9</sub> GlcNAc <sub>2</sub>	$0.23 \pm 0.01$	0.1

Tableau 5.1 – Affinité du CRD de la Langérine pour des monosaccharides. [92]

Sous une forme oligomérique, la présentation de 3 CRDs induit un phénomène d'avidité conduisant à des affinités apparentes augmentées. Il est donc plus aisé d'observer des interactions avec le domaine extracellulaire complet de la protéine.

### 5.2.2.1 Reconnaissance des glucides de type mannose, N-acetyl-glucosamine et fucose

On s'attend à ce que la Langérine reconnaisse tous les saccharides possédant des mannoses, N-acetyl-glucosamines ou des fucoses avec les hydroxyles 3 et 4 libres. Or, parmi tous les oligosaccharides possédant ces unités, la Langérine n'en reconnaît qu'une partie.

En effet, il a été montré que la Langérine reconnaît des structures saccharidiques possédant des mannoses terminaux ainsi que des structures composées d'un fucose terminal lié à un galactose [99]. Puis Feinberg et al. ont montré une forte spécificité de reconnaissance pour les oligosaccharides de type « High-Mannose » (Figure 5.8) très présents à la surface de glycoprotéines virales [38]. Ces résultats montrent aussi la capacité d'interaction de la Langérine avec des sucres terminant par un fucose ou un N-acetyl-glucosamine. Mais les liaisons saccharidiques impliquées révèlent des spécificités pour chaque lectine. En effet, seuls les oligosaccharides terminant par GlcNac $\alpha$ 1-2 (ou 1-3) Gal ont été reconnus ainsi que ceux terminant par un fucose lié à un galactose.

### 5.2.2.2 Reconnaissance des saccharides sulfatés

La première étude de spécificité de la Langérine a été effectuée par un criblage sur une puce à sucre par Galustian et al. en 2004 [44]. Ce premier travail a montré une faible interaction de la Langérine pour les saccharides de type « High Mannose » mais une forte interaction avec le Lewis x exposant un galactose terminal sulfaté en position 6. Cette spécificité de reconnaissance est troublante car la Langérine possède le motif EPN, qui permet une spécificité pour les saccharides de type mannose, et non QPD, qui procure une spécificité pour le galactose. La faible interaction observée est probablement due à la faible sensibilité de l'expérience. Tateno et al, ainsi que Feinberg et al confirment la spécificité de reconnaissance du galactose sulfaté en position 6 (Figure 5.8) [38, 99]. Ils observent aussi la reconnaissance de N-acétyl-glucosamine sulfaté en position 6. Cette particularité d'interaction est au coeur du travail de thèse effectué et en particulier les sucres de type GlcNac sulfaté.

### 5.2.2.3 Reconnaissance des saccharides bactériens et de champignons

D'autres résultats avec une puce comportant des motifs saccharidiques bactériens ont été obtenus (Figure 5.9). Ces résultats sont les premiers à suggérer une reconnaissance d'un motif interne à l'oligosaccharide. En effet, ils proposent que l'interaction est effectuée par le motif interne Man $\alpha$ 1-2Man[39].

Effectivement, la Langérine interagit fortement avec les polysaccharides portés par des souches d'*Escherichia coli* ou même de *Shigella boydii*. Elle reconnaît aussi les  $\beta$ -



glucans portés à la surface de levure comme par exemple les souches de *Malassezia furfur*, *Saccharomyces* et *Candida* (Figure 2.10)[23].

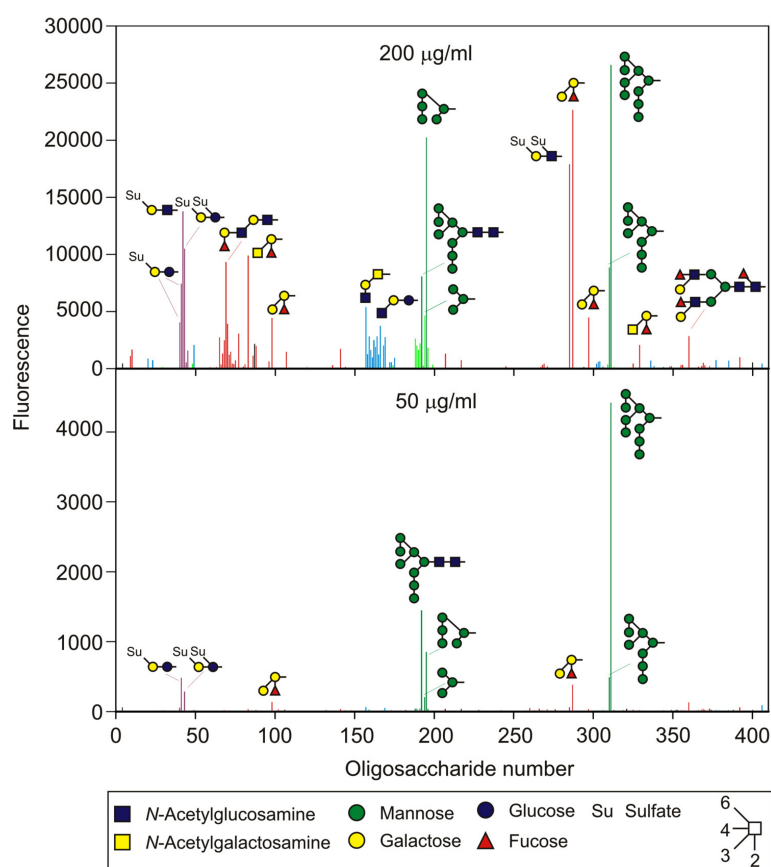


Figure 5.8 – Interaction de Langérine fluorescente avec une puce à glycanes. Les barres vertes représentent les structures oligosaccharidiques ou polysaccharidiques de type « High-Mannose ». Les barres bleues sont les oligosaccharides terminant par un N-acetylglucosamine. Les rouges sont les oligosaccharides terminant par un fucose. Les barres violettes représentent les oligosaccharides terminant par un galactose sulfaté en position 6 [38].

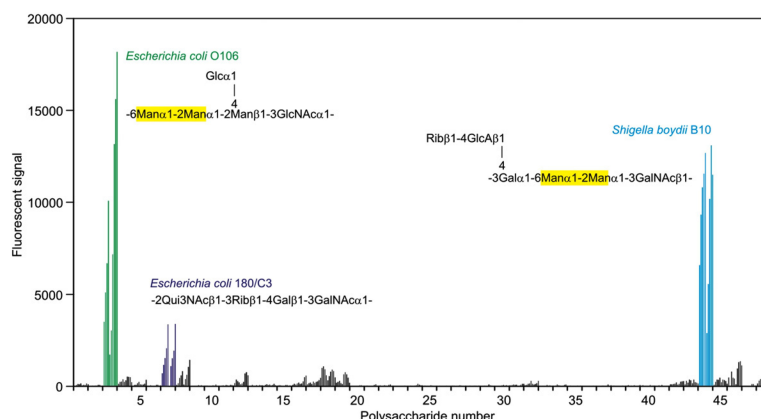


Figure 5.9 – Interaction de la Langérine avec une puce fonctionnalisée de polysaccharides bactériens [39].

Pour résumer la spécificité de reconnaissance de la Langérine (Figure 5.10) : cette lectine interagit principalement avec des structures saccharidiques possédant des mannoses ou N-acétyl-glucosamine terminaux, des fucoses terminaux liés à un galactose et aussi des saccharides sulfatés de type galactose ou N-acétyl-glucosamine sulfaté en position 6 [53]. Comparée à DC-SIGN, lectine très proche de la Langérine, la Langérine n'est pas capable de reconnaître toutes les structures possédant des mannoses et encore moins les oligosaccharides terminant par un fucose. Par contre, elle se démarque par sa capacité d'interaction avec des structures saccharidiques sulfatées.

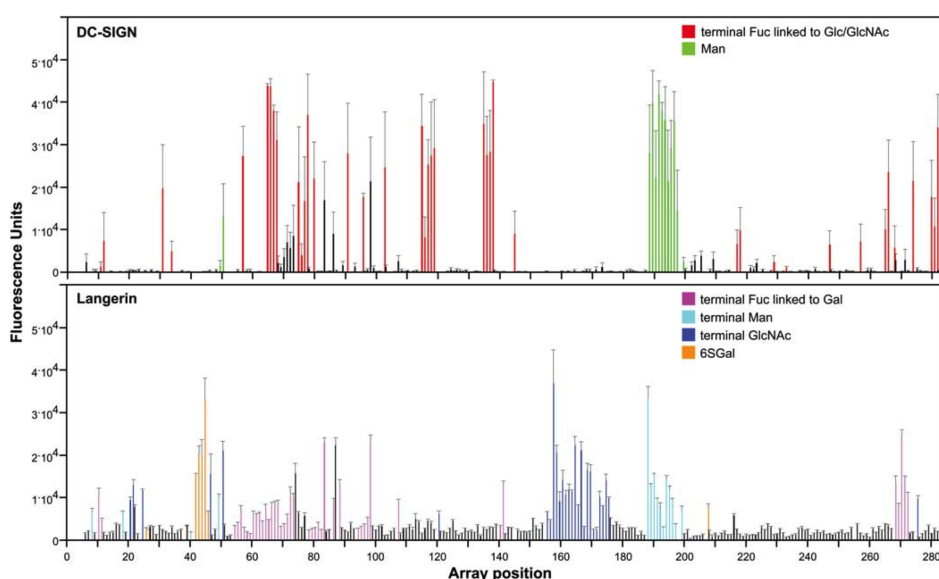


Figure 5.10 – Spécificité de reconnaissance de DC-SIGN (haut) et la Langérine (bas). [53]

### 5.2.3 Mode d'interaction de la Langérine

#### 5.2.3.1 Interaction avec les sucres sulfatés.

L'interaction particulière de la Langérine avec des sucres sulfatés se trouve au coeur de l'étude présentée dans cette thèse. L'observation de la structure cristallographique du CRD de la Langérine, et sa comparaison avec d'autres lectines du même groupe, montre effectivement la présence de deux résidus lysines proches du site calcium (Figure 5.11 a). Ceux-ci, en interaction avec le sulfate en position C6, permettent de stabiliser une orientation du galactose qui se retrouve capable d'interagir avec le calcium, par son groupement hydroxyle en position C3, et avec les résidus Asn307 et Glu293 par le groupement hydroxyle en position C4 (Figure 5.11, c). Cette position est inhabituelle. Les structures disponibles en présence de mannose (ou d'oligomannose) montrent une orientation différente du sucre en interaction avec le calcium (Figure 5.11 d) mais l'orientation du galactose mime celle du mannose pour les hydroxyles C3 et C4 ainsi le carbone 2 est en interaction avec l'alanine 289 (Figure 5.11, e). Ainsi, le groupement sulfate du galactose est fortement stabilisé par un réseau d'interaction avec la Lys313, la Phe315, l'Asn297 et la Lys299 (Figure 5.11, f). L'interaction ionique entre les deux résidus lysines et le groupement sulfate compense l'interaction qui n'est pas optimale avec le galactose dans un site possédant le motif EPN.

#### 5.2.3.2 Interaction avec le mannose

Dans le cas de l'interaction avec le mannose, celui-ci peut adopter deux orientations différentes mais avec des similitudes (Figure 5.12, A et B). En effet, on retrouve deux orientations qui interagissent de la même façon avec l'ion calcium mais avec une rotation de 180°. Ces deux orientations diffèrent dans les liaisons annexes et plus particulièrement avec le résidu Lys299 (Figure 5.12, B-C-D). Effectivement, dans la première orientation (bleu), la Lys299 interagit avec l'hydroxyle anomérique alors que dans la seconde (jaune) c'est l'hydroxyle en C2 qui est en interaction. Pour un disaccharide Man $\alpha$ 1-2Man on retrouve la double orientation soit avec le premier saccharide en interaction avec le calcium, soit avec le second (Figure 5.12, C-D).

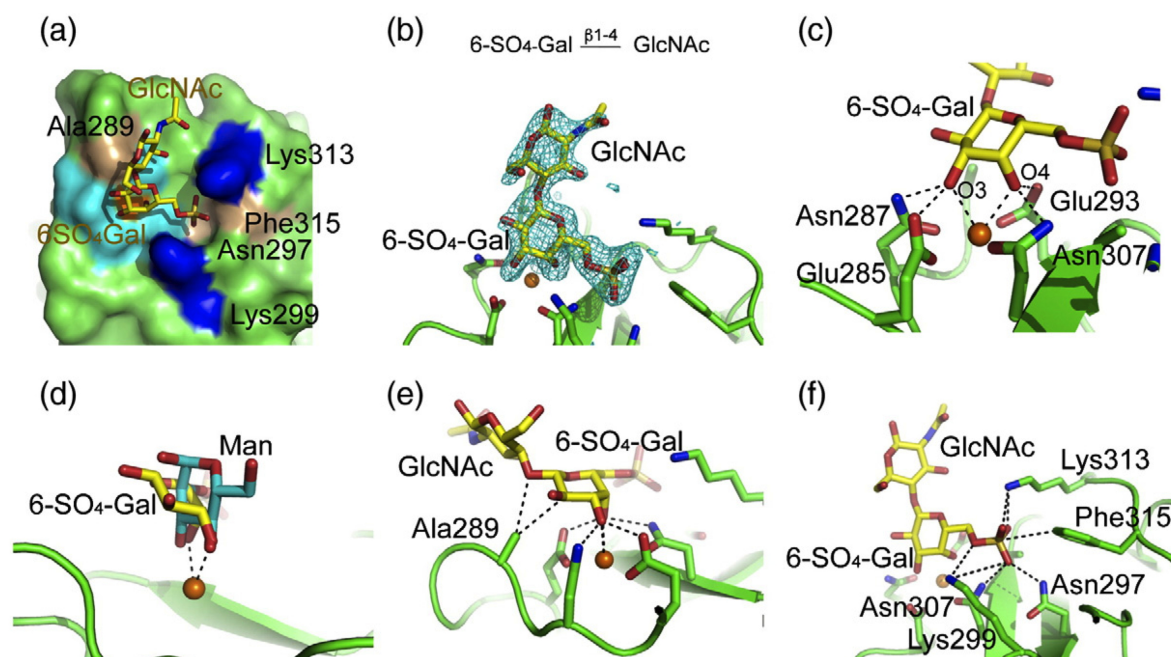


Figure 5.11 – Structure cristallographique du CRD de la Langérine en présence d'un 6 sulfogalactose $\beta$ 1-4 N-acetylglucosamine. a : Vue de la surface de la Langérine en interaction avec le disaccharide. b : Zoom sur le mode d'interaction du disaccharide et de sa densité électronique avec la Langérine. c : Zoom de l'interaction du galactose sulfaté avec l'ion calcium. d : comparaison de l'orientation du galactose sulfaté et du mannose en interaction avec l'ion calcium. e : Implication de l'hydroxyle en C2 et de la liaison osidique dans l'interaction. f : Zoom sur le réseau d'interaction engendré par le sulfate en position 6 du galactose. [39]

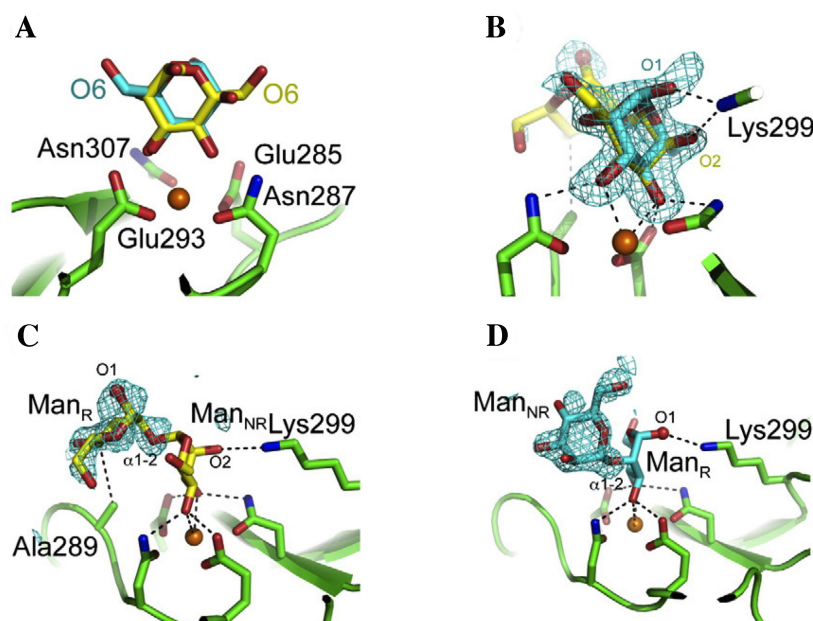


Figure 5.12 – Interaction du CRD de la Langérine avec un disaccharide  $\text{Man}\alpha 1\text{-}2\text{Man}$ . A : Orientation mettant en évidence les deux orientations possibles de mannose en interaction tout en gardant une superposition parfaite des hydroxyles en C3 et C4. B : Implication de la lysine 299 dans les deux orientations, soit en interaction avec l’hydroxyle anomérique soit avec celui en C2. C : Mode d’interaction avec le  $\text{Man}\alpha 1\text{-}2\text{Man}$  lorsque le mannose de l’extrémité non réductrice est en interaction avec le calcium. D : Mode d’interaction avec le  $\text{Man}\alpha 1\text{-}2\text{Man}$  lorsque le mannose de l’extrémité réductrice est en interaction avec le calcium.[39]

La spécificité de reconnaissance des structures saccharidiques possédant un mannose terminal est probablement due à l’implication du résidu lysine 299 dans le mode d’interaction et de l’encombrement apporté au niveau du site calcium. En effet, dans les modes d’interactions entre la Langérine et des oligosaccharides de type « High Mannose », seul le mannose terminal est capable d’interagir avec la Langérine car les autres utilisent pour la liaison osidique soit les hydroxyles anomériques et celui en C2 simultanément soit l’hydroxyle en C3 qui est indispensable pour l’interaction avec le calcium.

### 5.2.3.3 Interaction avec le fucose

La particularité de la reconnaissance du fucose par la Langérine est le fait que le saccharide doit être lié avec un galactose. La structure cristallographique du CTLD de la Langérine en présence du trisaccharide du groupe sanguin B montre l’importance de l’orientation des saccharides (Figure 5.13). Tout d’abord, le fucose interagit avec le calcium et les résidus coordinants par les hydroxyles en C2 et C3, ceci permet une liaison hydrogène entre la lysine 299 et l’hydroxyle en C4 (Figure 5.13, b-d). L’orientation du fucose permet de positionner le galactose terminal de telle façon qu’un réseau d’interaction

est effectué entre les hydroxyles en C3, C4 et C6 et les résidues 282, 283, 284, 285 et 287 de la Langérine. Lorsque l'on aligne les structures du complexe DC-SIGN/Lewis x et de la Langérine/trisaccharide du groupe sanguin B en orientant le fucose comme dans la structure de la Langérine (Figure 5.13, g), on observe un clash entre le galactose terminal du lewis x et de la Langérine. Ceci est dû à l'orientation imposée par la liaison fucose/N-acetyl-glucosamine. En effet, le galactose du lewis x étant lié au N-acetyl-glucosamine en  $\alpha$ 1-4, il se retrouve orienté en direction du squelette protéique. Par contre, si on aligne les structures avec le fucose interagissant dans le site calcium avec les hydroxyles en C3 et C4, comme dans la structure de DC-SIGN, les saccharides suivants ne sont pas en clash stérique avec la protéine mais aucune autre interaction permettant la stabilisation est possible par la Langérine. Ceci est dû à l'absence des résidus Asp367, Lys368, Leu371 et Lys373 remplacés par l'Ala309, Pro310, Lys313 et la Phe315 dans la Langérine.

Comme nous avons pu le voir, le motif permettant la coordination du calcium procure une première spécificité des monosaccharides pouvant être reconnus mais l'environnement du site impose des spécificités supplémentaires.

Donc la Langérine est un récepteur d'agents étrangers portant à leur surface des structures saccharidiques de type « High-Mannose » ainsi que le motif interne  $\text{Man}\alpha 1\text{-2Man}$  présent au niveau des lipopolysaccharides bactériens et les  $\beta$ -glucanes. La présence des lysines 299 et 313 attribue une capacité supplémentaire d'interaction avec des sucres sulfatés et notamment le galactose non reconnu sans sulfatation.

Parmi les différents ligands reconnus par la Langérine, les structures oligasaccharidiques contenant des GlcNac et GlcNac6S ont été très peu étudiées. Les modes d'interaction avec ce type de structure glucidique ont été donc étudiés dans ce travail de thèse (cf chapitre 13).

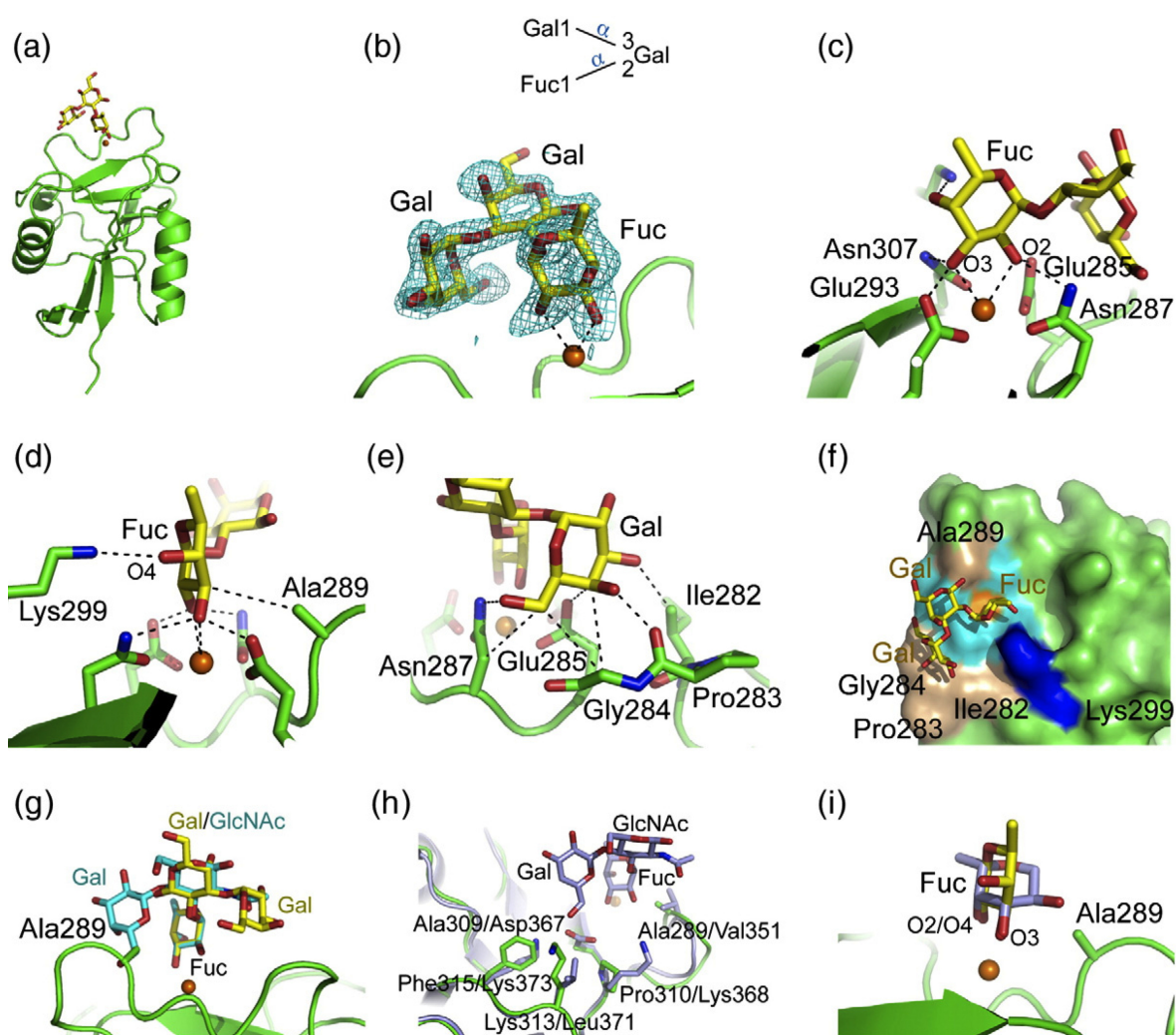


Figure 5.13 – Interaction du CTLD de la Langérine avec le trisaccharide du groupe sanguin B. a-f : différentes orientations de l'interaction de la Langérine avec le trisaccharide du groupe sanguin B. g : superposition du lewis x et du trisaccharide avec le mode d'interaction de la Langérine. h : superposition du lewis x et du trisaccharide avec le mode d'interaction de DC-SIGN. i : orientation du fucose en interaction via les hydroxyles en C2 et C3 (jaune) ou avec les hydroxyles en C3 et C4 (mauve). [39]

---



# Chapitre 6

## Les hydrates de carbone étudiés et leurs fonctions

### 6.1 Les hydrates de carbone de type « High Mannose »

Les protéines d'enveloppe de certains virus, comme le VIH (la dengue, ebola), sont riches en oligosaccharide de type « High Mannose ». Elles forment un nuage autour des protéines et permettent aux cellules de reconnaître le virus. Comme évoqué précédemment, les virus peuvent aussi utiliser cette capacité de reconnaissance pour infecter la cellule.

Ces structures oligosaccharidiques sont des N-glycosylations effectuées post-traductionnelle sur des asparigines (Figure 6.1). On retrouve différentes structures dépendantes du nombre de mannoses présents. Celui-ci peut aller de 5 à 9. Tous les mannoses sont liés par une liaison  $\alpha$  soit de type 1-2, 1-3 ou 1-6. Les mannoses terminaux, reconnus par la Langérine (chapitre 5), sont tous présentés par une liaison  $\alpha$ 1-2. Ce type de structure permet aussi de créer un phénomène d'avidité par la présentation multiple du ligand. Comme on peut le voir sur la figure 6.3, trois mannoses terminaux sont présentés, ceci peut théoriquement interagir avec plusieurs CRDs d'une lectine dans la limite des encombrements stériques occasionnés. Au-delà de cette présentation de plusieurs sucres terminaux, les glycoprotéines possèdent souvent de nombreux sites de glycosylation ce qui augmente d'autant plus l'effet d'avidité.

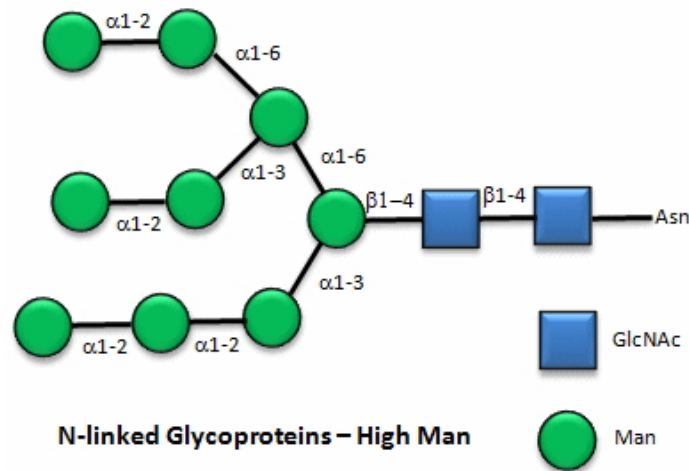


Figure 6.1 – Schéma représentatif des structures de type « High-Mannose ».

## 6.2 Les glycosaminoglycanes (GAGs)

Les glycosaminoglycanes sont impliqués dans de nombreux processus physiologiques et pathologiques [11, 57]. Ils permettent la régulation de facteurs de croissance, des chimiokines et cytokines [57]. Ces sucres sont très présents dans les tissus conjonctifs et la matrice extracellulaire.

Les glycosaminoglycanes forment une famille de sucres particuliers pour plusieurs raisons : tout d'abord ces sucres possèdent un motif disaccharidique répété. Ensuite ils sont liés à des protéines pour former le protéoglycane (O-glycosylation, sauf l'acide hyaluronique). Et enfin ce sont des sucres sulfatés (sauf l'acide hyaluronique).

Ils sont composés d'un disaccharide contenant une hexosamine (glucosamine ou galactosamine) et d'un autre ose (de type acide glucuronique, acide iduronique ou d'un galactose) (Figure 6.2). La glucosamine est soit N-acétylée soit N-sulfatée alors que la galactosamine est toujours N-acétylée. Ce sont des chaînes linéaires de haut poids moléculaires (de 10 à 100 kDa) (Figure 6.2).

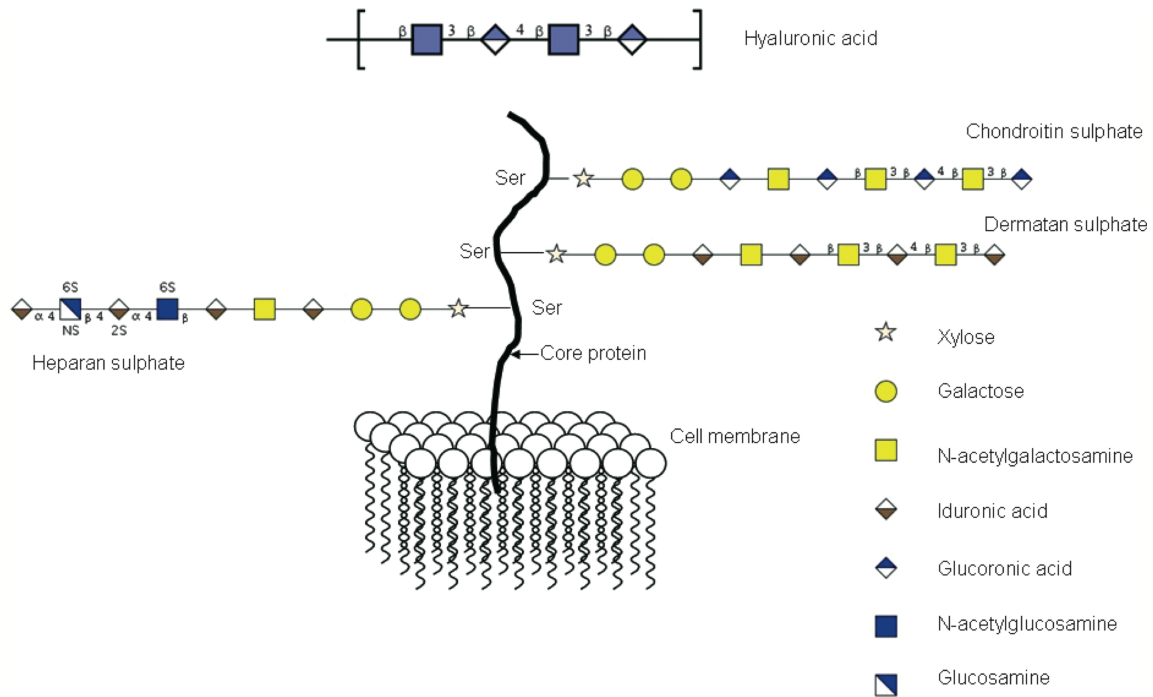


Figure 6.2 – Représentation schématique des différents glycosaminoglycanes retrouvés. [45]

On peut séparer cette famille en plusieurs groupes :

- Les chondroitines sulfates (introduisant aussi le dermatane sulfate)
- Les héparanes sulfates (HS) et héparines
- Le kératane sulfate
- L'acide hyaluronique (non discuté ici)

Les chondroitines sulfates sont composées d'un dissaccharide formé par un acide glucuronique lié en  $\beta 1-3$  à un N-acétyl-galactosamine, ce motif est répété avec une liaison  $\beta 1-4$ . On retrouve le long de la chaîne différents motifs de sulfatation (Figure 6.3) :

- Le motif A présentant une sulfatation en position 4 de la galactosamine.
- Le motif B (ou dermatane sulfate) composé plutôt d'un acide iduronique sulfaté en position 2, à la place de l'acide glucuronique, et d'un galactosamine sulfaté en position 6.
- Le motif C présentant une sulfation en position 6 de la galactosamine.

Le kératane sulfate est lui composé d'un galactose plus ou moins sulfaté en position 6 lié en  $\beta 1-4$  à un glucosamine sulfaté ou non en 6. Ceci est répété par une liaison  $\alpha 1-3$ . La chaîne polysaccharidique du kératane sulfate se finit par un acide sialique (sucre important dans la reconnaissance du soi).

## 6.2. LES GLYCOSAMINOGLYCANES (GAGS)

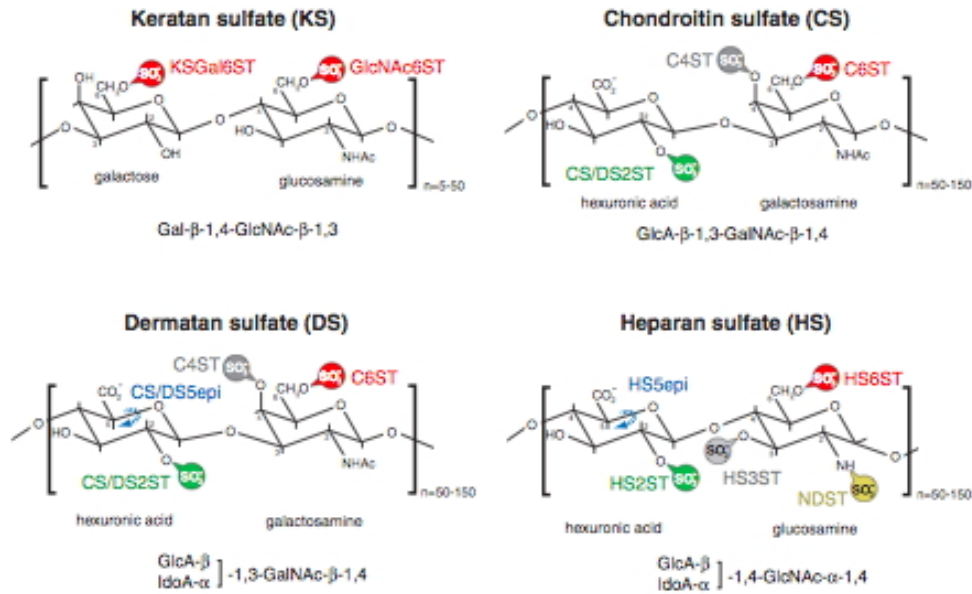


Figure 6.3 – Illustration de la composition des différents groupes de glycosaminoglycanes. Les différentes modifications possibles sont représentées en rouge, vert et jaune pour les différentes sulfotransférases et en bleu pour les épimérases. Le dermatane sulfate correspond à la chondroïtine sulfate B. [9]

Enfin, le groupe des héparanes sulfates (HS) est lui composé soit d'un acide glucuronique soit d'un acide iduronique sulfaté en position 2 lié en  $\beta$ 1-4 à un glucosamine plus ou moins sulfaté en position N et 6. Le motif est répété par une liaison  $\alpha$ 1-4. Mais le long d'une chaîne d'héparane sulfate, on retrouve différents motifs (Figure 6.4). Un motif fortement sulfaté nommé NS, un autre faiblement sulfaté appelé NA/NS et enfin un dernier non sulfaté dénommé NA. Cette hétérogénéité de la chaîne d'héparane sulfate permet donc des reconnaissances différentes selon le domaine utilisé dans l'interaction.

Les HS jouent aussi un rôle dans l'infection par le VIH. La glycoprotéine d'enveloppe du VIH, gp120, peut interagir avec les HS permettant de concentrer le virus à proximité des récepteurs d'entrée ainsi que de protéger le virus à l'état infectieux pendant plusieurs jours [6]. Ces HS interviennent dans le passage du virus à travers l'épithélium des muqueuses mais seuls les virus de tropismes X4 sont capables d'interagir avec les HS par la présence de résidus chargés positivement dans la boucle V3 de gp120 [7][73]. Ainsi d'autres virus comme l'herpès simplex utilisent les HS portés par les protéines syndecans pour l'attachement du virus à la surface cellulaire [2].

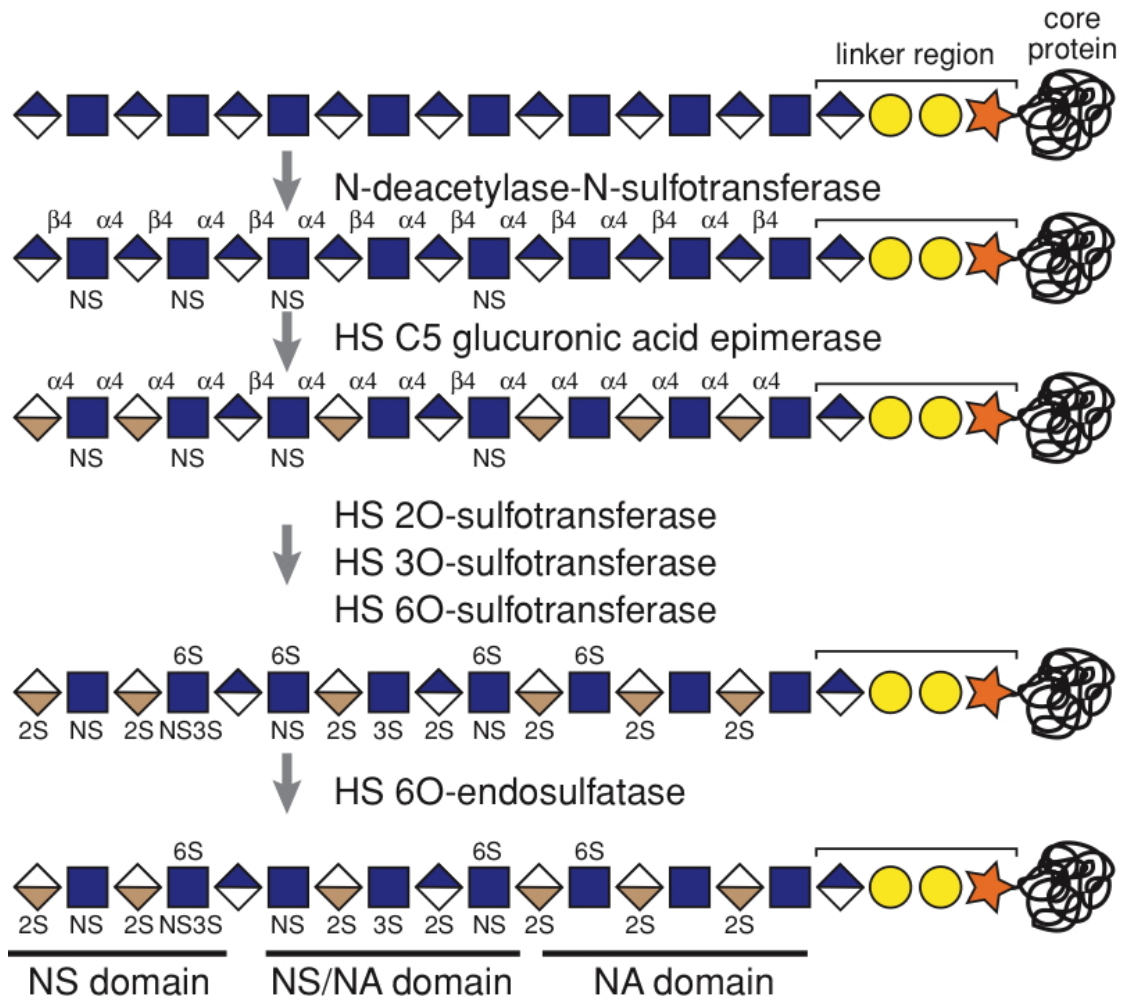


Figure 6.4 – Représentation des différentes modifications pouvant être effectuées sur les chaînes d'héparane sulfate révélant la diversité des structures saccharidiques. Les modifications sont réalisées par des épimérase et des sulfotransférases dans l'appareil de Golgi. [9]

La Langérine a été décrite comme interagissant avec les kératanes sulfates [99]. Par cette observation couplée à l'interaction décrite pour certains sucres sulfatés, cela nous a conduit à évaluer la capacité de la Langérine à reconnaître les GAGs.

# Deuxième partie

## Les axes de travail

---

La Langérine a tout d’abord été identifiée comme une lectine de type-C spécifique des cellules de Langerhans induisant la formation des granules de Birbeck[104]. Comme évoqué précédemment, la seule transfection du gène de la Langérine dans des fibroblastes murins induit la formation d’organite cellulaire de structure similaire au granule de Birbeck des cellules de Langerhans. Par le séquençage de cette protéine, Valladeau et al ont pu mettre en évidence l’appartenance de la Langérine à la famille des lectines de type-C par la présence, entre autres, du CRD contenant le motif EPN [104]. Par la suite, les études de la spécificité de reconnaissance de la Langérine ont permis de mettre en évidence la capacité d’interaction de la Langérine avec des sucres sulfatés.

Donc, deux axes majeurs ont été fixés pour cette thèse autour de la caractérisation structurale et fonctionnelle de la Langérine :

1. Dans une première partie du travail, nous avons voulu caractériser l’implication de la Langérine dans la formation des granules de Birbeck. Ceci se base sur les études menées par Jenny Valladeau et Pauline Verdjik qui ont montré qu’une simple mutation ponctuelle de la Langérine pouvait dramatiquement affecter la formation et la structure du granule de Birbeck (Figure 7.1)[105, 107]. De plus, la mutation W264R étant une forme polymorphique humaine, l’identification des conséquences sur le fonctionnement de la Langérine peut sembler utile du fait de l’implication de cette protéine vis-à-vis du VIH. Deux simples mutations dans le domaine de reconnaissance des hydrates de carbones conduisent pour l’une à la perte des granules de Birbeck (mutation W264R, Figure 6.7, C), et pour l’autre à une modification radicale de la structure (mutation F244L, Figure 6.7, B). Par la localisation des mutations dans le CRD, et la nécessité de sa présence pour la structure du granule[100], l’étude de ces mutations pourrait apporter des informations sur la formation des BGs. Des études de biologie cellulaire, de biochimie et de biologie structurale ont été menées tout d’abord pour vérifier l’effet des mutations sur la protéine humaine à l’échelle cellulaire pour ensuite se focaliser sur les effets intra-moléculaires de ces mutations sur la Langérine.
2. Le deuxième axe est basé sur la capacité particulière de la Langérine à se fixer à des sucres sulfatés. Cette capacité est encore très peu étudiée et peut avoir un rôle physiologique important. Parmi les ligands naturels possédant ce type de saccharide sulfaté, on peut penser aux glycosaminoglycanes. Les glycosaminoglycanes se révèlent être très présents dans la matrice extracellulaire ainsi que la Langérine. Nous avons donc choisi d’étudier les possibles interactions entre ces deux composants.

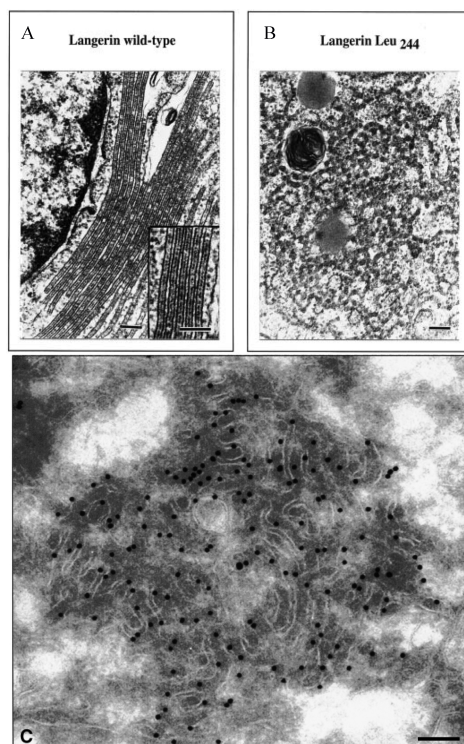


Figure 6.5 – Modification de la structure du granule de Birbeck par la mutation F244L et perte de la formation du granule par la mutation W264R. A : fibroblaste transfecté avec le gène de la Langérine de souris. B : Fibroblaste transfecté avec le gène de la Langérine de souris contenant la mutation F244L[105]. C : Cellule de Langerhans humaine contenant le gène de la Langérine avec la mutation W264R [107].

3. En marge des deux questions précédentes qui représentent les travaux principaux de mon projet de thèse, je me suis aussi impliqué dans un projet de développement d'inhibiteur de DC-SIGN au sein d'un projet européen "CARMUSYS". Ce projet, satellite pour moi, a pour but de développer des molécules inhibitrices de DC-SIGN afin d'empêcher les cellules dendritiques de promouvoir l'infection par le VIH. Pour autant, l'objectif est de produire des molécules spécifiques de DC-SIGN et non de la Langérine afin de préserver son effet protecteur.

Dans ce contexte, je me suis donc attelé à produire différentes formes recombinantes de la Langérine afin de pouvoir mener une étude fonctionnelle et structurale sur cette lectine. Pour ces études, différentes techniques ont été utilisées comme l'observation de l'évolution de la fluorescence (« thermal shift assay »), le dichroïsme circulaire, la microscopie électronique, la résonance plasmonique de surface et la cristallographie des rayons X. De plus, des collaborations avec des équipes de biologie cellulaire et de RMN ont été menées pour étayer les résultats obtenus.



# Troisième partie

## Matériel et méthodes

---

# Chapitre 7

## Biologie moléculaire

Dans cette partie je ne développerai que les aspects expérimentaux qui ne sont pas traités dans les articles.

### 7.1 Constructions plasmidiques utilisées

Les constructions initiales utilisées ont été réalisées avant mon arrivée. Nous exprimons les protéines de façon recombinante chez les bactéries. Pour cela tous les gènes codant pour les protéines étudiées dans cette thèse ont été clonés dans le vecteur pET30b (Novagen) (Figure 8.1). L'expression des gènes clonés dans le site de clonages multiples (PCS) dans le vecteur pET30b est régulée par un promoteur de phage T7 inhibé par le represseur lac. Cette inhibition est levée par l'ajout d'isopropyl- $\beta$ -D-thiogalactoside (IPTG). L'ajout d'IPTG permet l'expression de la protéine clonée dans le PCS.

La première construction utilisée, pET-30b Lg-S-CRD, correspond à l'insertion dans le vecteur pET-30b de la région codant pour le domaine CRD de la Langerine humaine clonée en aval du peptide d'adressage périplasmique OmpA et d'une étiquette de type « StrepTag II » (Annexe 1). Le clonage de la région codant pour le domaine CRD en aval du peptide d'adressage périplasmique OmpA permet de localiser les néo-protéines exprimées dans un espace cellulaire oxydant (le périplasme) permettant de former les ponts disulfure essentiels à la structuration du domaine CRD et donc de l'obtenir de façon soluble après la casse des bactéries. La séquence d'adressage périplasmique OmpA ainsi que l'étiquette « StrepTag II » ont été ajoutée par le clonage dans le vecteur pASK-IBA6 (IBA). Par la suite l'ensemble a été sous-cloné dans le vecteur pET-30b apportant la résistance à l'antibiotique Kanamycine (Figure 7.1). Le clonage a été réalisé dans le PCS entre les sites de restrictions *HindIII* et *BamHI*.

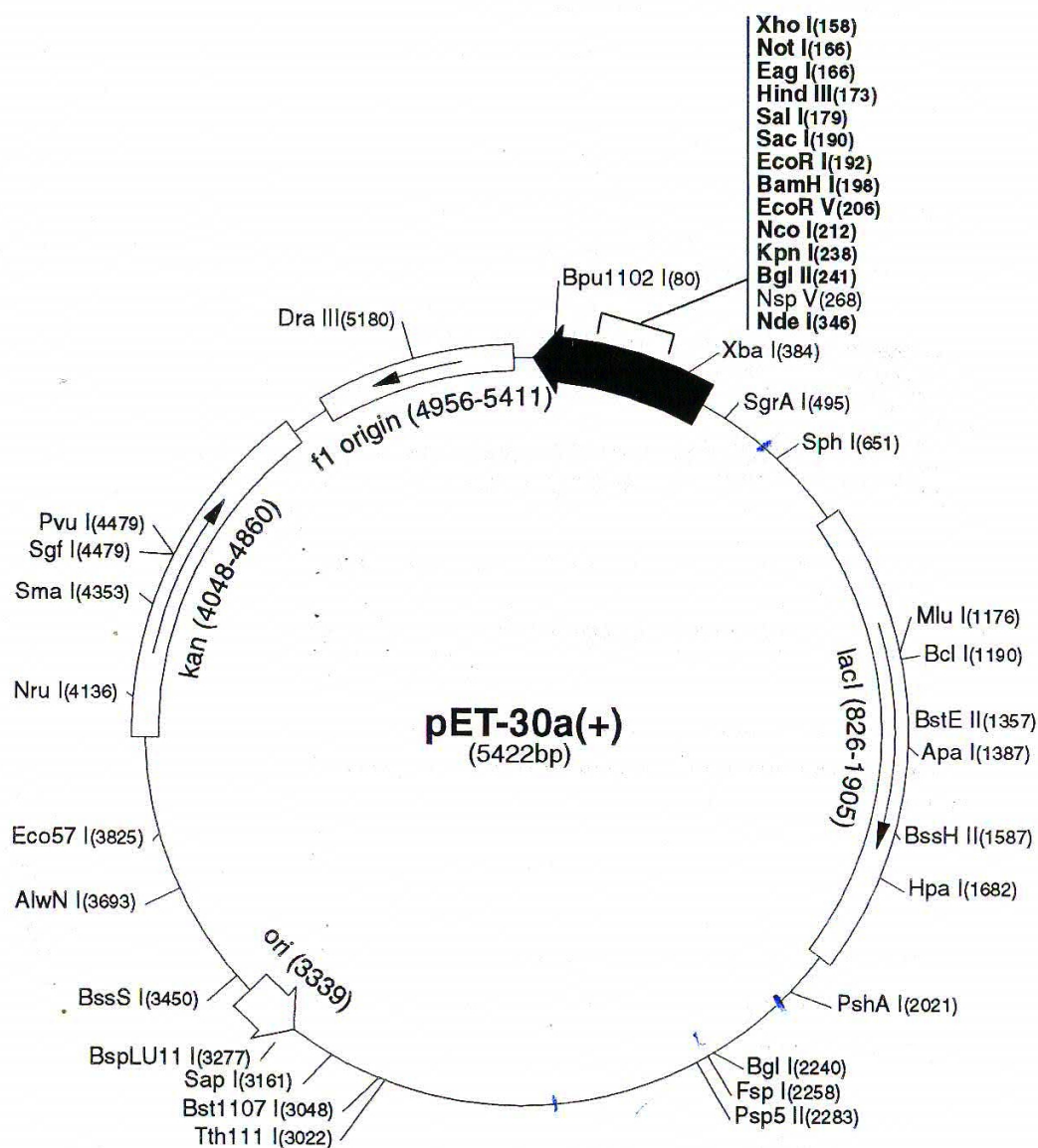


Figure 7.1 – Carte du plasmide pET30b.

La seconde construction, pET-30b Lg-ECD, correspond à l'insertion dans le vecteur pET-30b de la région codant pour le domaine extracellulaire de la Langerine humaine (Annexe 2).

Enfin la troisième construction utilisée, pET-30b Lg-S-ECD, correspond à l'insertion dans le vecteur pET-30b de la région codant pour le domaine extracellulaire clonée en aval d'une étiquette de type « StrepTag II » (Annexe 3).

## 7.2 Mutagenèse dirigée de Lg-S-CRD et Lg-ECD

Afin d'étudier l'influence de résidus dans la structure et la fonction des protéines, nous avons effectué des mutants de ces protéines en mutant les codons des résidus d'intérêt par mutagenèse dirigée.

Mutation	Amorce
K299A	5'-AATGAACACTGTGGCAATATAGCGGCTCCCTCACTTCAGGCCTGG-3'
	5'-CCAGGCCTGAAGTGAGGGAGCCGCTATATTGCCACAGTGTTCATT-3'
K313A	5'-TGGAATGATGCCCCATGTGACGCAACGTTTCTTTTCATTGTAAAG-3'
	5'-CTTACAAATGAAAAGAAACGTTGCGTCACATGGGGCATCATTCCA-3'
K257A	5'-TGGATTGGCCTGACTGCCGCAGGGATGGAAGGG-3'
	5'-CCCTTCCATCCCTGCGGCAGTCAGGCCAATCCA-3'
N297A	5'-GGGAACAATGAACACTGTGGCGCTATAAAGGCTCCCTCACTTCAG-3'
	5'-CTGAAGTGAGGGAGCCTTTATAGCGCCACAGTGTTCATTGTCCC-3'

Tableau 7.1 – Liste des oligonucléotides utilisés pour produire les mutants de la Langérine.

L'amplification des ADNs a été réalisée par Polymerase Chain Reaction (PCR) en utilisant le mélange et le protocole suivant :

– Milieu réactionnel :

1,25 µL de chaque amorce à 100 ng/µL

1 µL de dNTPs à 10 mM

1 µL d'ADN matrice à 50 ng/µL

5 µL de tampon 10x

1 µL de polymérase Pfu Turbo à 2,5 Unités/µL

qsp 50 µL H<sub>2</sub>O

– Programme de PCR : 1) 30 s à 95°C ; 2) 30 s à 95°C ; 3) 1 min à 55°C ; 4) x min à 68°C ; 5) 18 retours à l'étape 2 ; 6) 4°C temps infini. La durée de l'étape 4 est de 4 min pour Lg CRD et de 15 min pour Lg ECD.

La matrice d'ADN est finalement digérée par ajout de 1 µL de Dpn1.

## 7.3 Transformation bactérienne

Les vecteurs contenant les séquences des protéines à étudier peuvent être ensuite introduits dans des bactéries soit pour amplifier et purifier les vecteurs (souche bactérienne utilisée : TOP10) afin de les séquencer et les stocker, soit pour l'expression des protéines (souche utilisée : BL21(DE3)).

Un aliquot de 200 µL d'*E. coli* est transformé avec 1 µL de vecteur contenant le gène codant pour la protéine à étudier. Le mélange est incubé pendant 1 h dans la glace. Un choc thermique d'une minute est ensuite réalisé en incubant le mélange dans un bain marie à 42°C. Puis on ajoute 1 mL de LB. Le tube est incubé pendant 1 h à 37°C dans un incubateur sous agitation. Le tube est centrifugé à 18 000 g pendant 10 min afin d'éliminer

### 7.3. TRANSFORMATION BACTÉRIENNE

---

le milieu. Le culot bactérien est resuspendu dans environ 100  $\mu\text{L}$  de LB avant d'étaler la solution sur une boîte de Pétri supplémentée en kanamycine (50  $\mu\text{g}/\text{mL}$ ). Les bactéries sont ensuite incubées à 37°C toute la nuit.

Les plasmides issus de mutagenèse dirigée ont été purifiés à partir de 5 ml de culture de bactéries Top10 transformées, à l'aide du kit Nucleospin Plasmid (Macherey-Nagel) et les produits de purification ont été vérifiés par séquençage (Génome Express).

# Chapitre 8

## Biochimie, résonance plasmonique de surface et microscopie électronique

### 8.1 Expression et purification des protéines

Plusieurs constructions de la Langérine ont été utilisées lors de cette thèse correspondant au domaine CRD précédé du « Strep Tag II » (Lg-S-CRD), du domaine ECD (Lg-ECD) et du domaine ECD précédé du « Strep Tag II » (Figure 8.1).

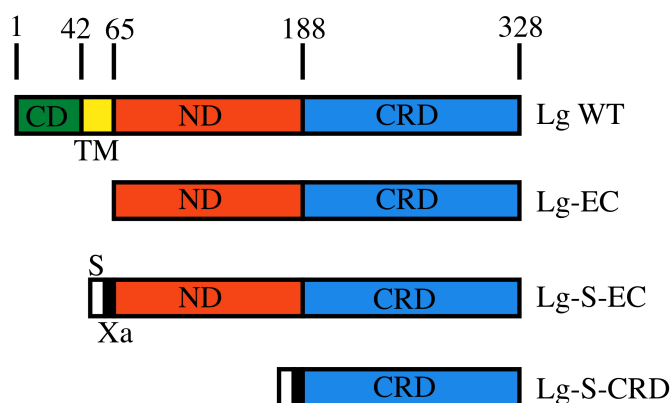


Figure 8.1 – Constructions protéiques obtenues.

#### 8.1.1 Expression et purification de Lg-S-CRD<sub>WT</sub>, Lg-S-CRD<sub>F241L</sub> et Lg-S-CRD<sub>K257A</sub>

Une pré-culture sur la nuit de 50 mL dans du milieu LB / kanamycine à 50 µg/mL réalisée à partir d'une colonie BL21 (DE3) isolée, transformée avec les différents vecteurs, permet d'ensemencer 1 L de milieu LB / kanamycine à 50 µg/mL. Lorsque la  $DO_{600nm}$

atteint une valeur de 2, sous agitation à 37°C, la surexpression protéique est induite avec 1 mM d'IPTG et le milieu est supplémenté avec du CaCl<sub>2</sub> 25 mM et laissé sur la nuit à 20°C. Dans chaque cas, des échantillons sont prélevés avant et après induction pour contrôler la surexpression par gel SDS-PAGE à 15%. Le culot bactérien de 1 L de culture obtenu après centrifugation (5000 g, 20 min, 4°C) est resuspendu dans 30 mL de tampon A (NaCl 150 mM; Tris-HCl 25 mM pH8; CaCl<sub>2</sub> 4 mM) supplémenté d'une pastille d'anti-protéases « Complete EDTA free » (Roche). Les cellules sont lysées par sonication (4 min de sonication à 45% d'amplitude, pulsation de 2 s avec une pause de 10 s entre chaque pulsation, Digital Sonifier® W-450 de Branson). Le lysat est centrifugé à 100 000 g, pendant 30 min, à 4°C. Le surnageant de centrifugation est purifié, à un débit de 3,5 mL/min, sur une colonne de streptactine de 5 mL (GE Healthcare), équilibrée en tampon A. La colonne est lavée avec 130 mL de tampon A et la protéine est éluée avec 50 mL de tampon A supplémenté avec de la D-Desthiobiotine 2,5 mM. L'élution des protéines est suivie par l'absorbance à 280 nm. Les fractions contenant la protéine d'intérêt sont ensuite concentrées à l'aide d'un concentrateur Vivaspın 20 (Vivascience) ayant un seuil de coupure de 10 000 Da. Une deuxième étape de purification est réalisée afin de séparer les protéines fonctionnelles. Pour cela, on utilise la capacité de reconnaissance du domaine CRD pour le mannose. La protéine concentrée est donc injectée sur une colonne de Mannose-Agarose équilibrée avec le tampon A. Les protéines présentant un retard à l'élution sont ensuite concentrées, à l'aide d'un concentrateur Vivaspın 20 (Vivascience) ayant un seuil de coupure de 10 000 Da, à une concentration de 5,5 mg/mL.

### 8.1.2 Expression et purification de Lg-ECD<sub>WT</sub>, Lg-ECD<sub>F241L</sub>, Lg-ECD<sub>K257A</sub>, Lg-ECD<sub>K299A</sub>, Lg-ECD<sub>K313A</sub>, Lg-ECD<sub>K299A K313A</sub>, Lg-S-ECD<sub>WTet</sub> Lg-S-CRD<sub>W264R</sub>

Le culot bactérien obtenu après centrifugation à 5 000 g de 1 L de culture pendant 20 min à 4°C est resuspendu dans 30 mL de tampon B ( NaCl 150 mM; Tris 25 mM pH 8) supplémenté d'une pastille d'antiprotéases « Complete EDTA free » (Roche). Après sonication, le lysat est centrifugé à 100 000 g, 30 min, à 4°C. Le culot contenant les corps d'inclusion est lavé par deux fois par une alternance de resuspension au potter et d'ultracentrifugation à 100 000 g, 30 min, à 4°C. Ces deux lavages sont réalisés successivement avec 30 mL de tampon B supplémenté avec 2 M Urée et 1% triton X-100 et 30 mL de tampon B. Le culot final des corps d'inclusion est ensuite solubilisé dans du tampon B supplémenté avec 6 M Guanidine et 0,01% de β-mercaptoéthanol à l'aide du potter. La fraction insoluble résiduelle est éliminée par ultracentrifugation. Le surnageant est dilué, par goutte à goutte, dans 120 mL de tampon NaCl 1,25 M; Tris-HCl 210 mM pH 8 et du



$\text{CaCl}_2$  25 mM. La suspension est centrifugée à 100 000 g, pendant 1 h, à 4°C. Le surnageant est dialysé contre 1 L d'eau sur la nuit, la dialyse est poursuivie contre du tampon A avec changement toutes les 4 h renouvelé 3 fois. Le produit de dialyse est centrifugé à 100 000 g, pendant 1 h, à 4°C. Le surnageant est ensuite purifié, à un débit de 2 mL/min, sur une colonne de Mannane-Agarose de 10 mL, équilibrée en tampon A. La colonne est lavée avec 150 mL de tampon A et les protéines sont éluées avec un tampon NaCl 150 mM ; Tris-HCl 25 mM pH 8 et de l'EDTA 1 mM. L'élution est suivie par l'absorbance à 280 nm. Les fractions contenant la protéine d'intérêt sont ensuite concentrées à l'aide d'un concentrateur Vivaspinn 20 (Vivascience) ayant un seuil de coupure de 10 000 Da. Les protéines sont dialysées contre du tampon A avec un facteur de dialyse de 100, le tampon est changé toutes les 4 h à 3 reprises.

Dans le cas de Lg-S-ECD<sub>WT</sub> et Lg-S-CRD<sub>W264R</sub>, une étape préliminaire de purification est réalisée par chromatographie d'affinité sur une colonne de streptactine en utilisant le même protocole que précédemment décrit pour les Lg-S-CRD. Ensuite les protéines sont concentrées à l'aide d'un concentrateur Vivaspinn 20 (Vivascience) ayant un seuil de coupure de 10 000 Da. Les Lg-S-CRD<sub>W264R</sub> sont séparées sur une colonne de Mannose-Agarose comme décrit pour les autres Lg-S-CRD. Dans le cas des Lg-S-ECD, les protéines sont purifiées en utilisant le même protocole que pour les autres formes de Lg-ECD.

## 8.2 La Résonance Plasmonique de Surface (SPR)

### 8.2.1 Principe

La résonance plasmonique de surface est un phénomène physique utilisé pour détecter des interactions entre un ligand et un récepteur à la surface d'une couche métallique (Figure 8.2).

En effet, à l'interface entre deux milieux d'indice de réfraction différents, un faisceau de lumière monochromatique polarisée arrive avec une incidence pour laquelle le faisceau lumineux est totalement réfléchi. Dans ces conditions une composante électromagnétique de la lumière appelée onde évanescente pénètre le milieu d'indice de réfraction inférieur sur une courte distance. L'interposition d'un film d'or à l'interface des deux milieux engendre une résonance entre les électrons délocalisés de la couche périphérique de l'or et l'onde évanescente, conduisant à une diminution de l'intensité lumineuse dans le faisceau réfléchi (signal de résonance plasmonique), et ceci dans un angle précis appelé angle de résonance. Cet angle de résonance dépend de plusieurs paramètres et notamment de l'indice de réfraction du milieu dans lequel l'onde évanescente se propage, c'est à dire le milieu aqueux.

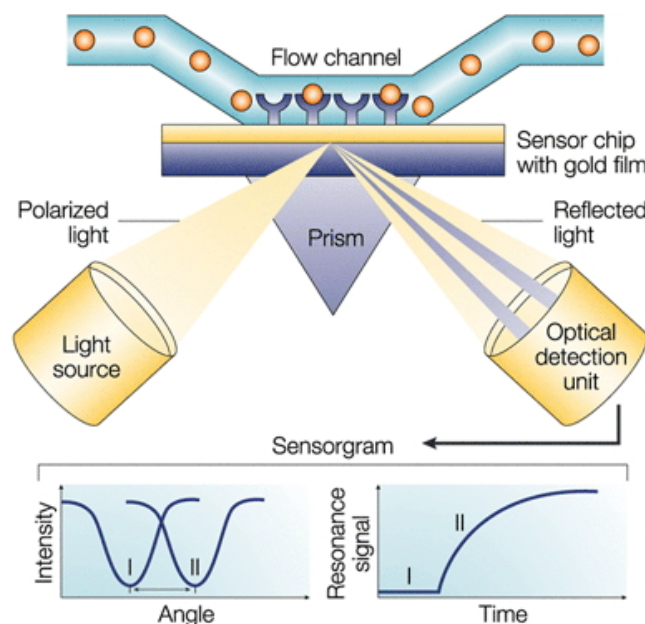


Figure 8.2 – Schéma de la technique de résonance plasmonique de surface, de la variation de l'angle de résonance et d'un sensorgramme montrant les différentes phases d'une interaction. [17]

Le système de détection permet d'observer en temps réel la variation de l'angle de résonance lorsque les biomolécules injectées interagissent avec celles immobilisées sur la surface. Ceci correspond au sensorgramme qui représente la variation de l'angle de résonance en unité de résonance (RU) en fonction du temps.

Les appareils les plus utilisés actuellement sont les biocapteurs Biacore. Ceux-ci sont composés de trois éléments principaux :

- La sensor chip : C'est une surface composée d'un support en verre (qui permet d'obtenir un indice de réfraction élevé conduisant à la propagation de l'onde évanescente dans le milieu aqueux) recouvert d'une fine couche d'or (permettant la résonance) et d'un hydrogel constitué d'un polymère de glucose fortement carboxyméthylé d'épaisseur variable. Cet hydrogel carboxyméthylé permet d'immobiliser de façon covalente une des deux molécules (appelée ligand) par un couplage chimique. Ce couplage est le plus souvent réalisé entre les groupes amines de la biomolécule et les carboxyles activés préalablement (Figure 8.3).
- Le système micro fluide : Ce système permet de contrôler le débit à la surface des canaux utilisées (de 1 à 100  $\mu\text{L}/\text{min}$ ) ainsi que l'injection de la molécule.
- Le faisceau lumineux monochromatique et polarisé : Un faisceau lumineux monochromatique et polarisé par un prisme illumine la sensor chip.

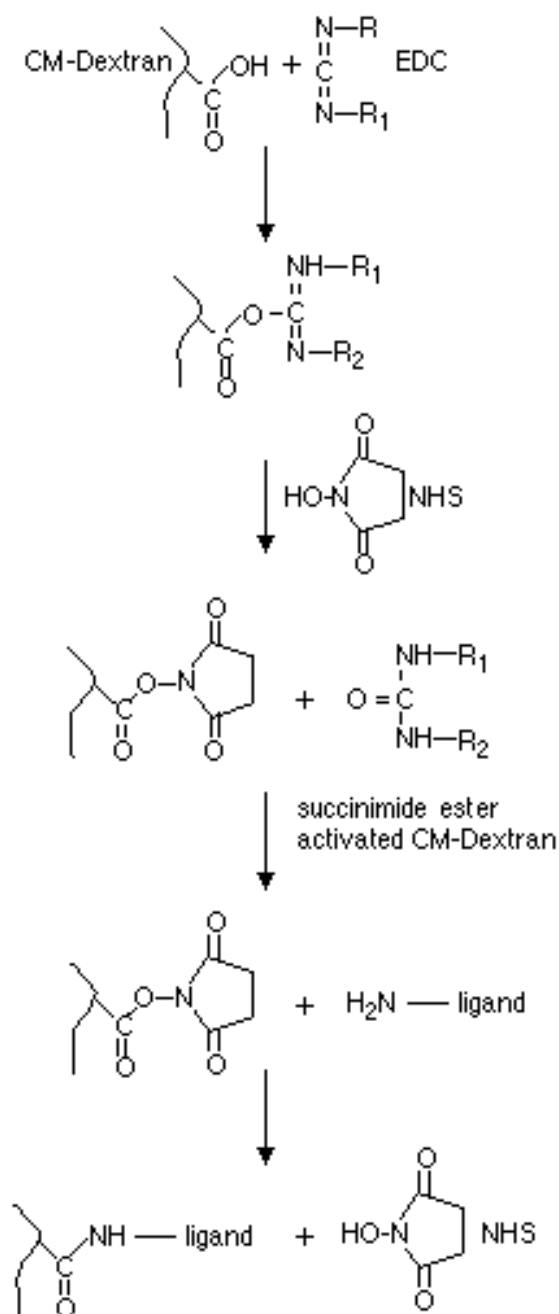


Figure 8.3 – Couplage covalent par les amines protonées d'une biomolécules. (note d'application GE Healthcare)

### 8.2.2 Fonctionnalisation des surfaces

L'appareil utilisé tout au long du travail de thèse est un Biacore 3000 qui permet d'utiliser quatres canaux en même temps et de disposer d'un robot d'injection.

Les puces utilisées sont des sensor chip CM4 présentant un polymère de carboxyl dextran.

La fonctionnalisation des surfaces est réalisée à un débit de 5  $\mu\text{L}/\text{min}$  dans un tampon commercial HBS-P (GE Healthcare). Les surfaces sont d'abord activées par l'injection de 50  $\mu\text{L}$  d'un mélange équimolaire d'EDC et de NHS (même volume et même concentration de 0,5 M) (Figure 8.3). Ensuite les protéines d'intérêt sont diluées dans un tampon composé de 5 mM d'acétate de sodium pH 4,5 afin d'être en dessous du point isoélectrique pour que les amines à la surfaces des protéines soient protonées et donc réactives. Après l'injection des protéines à immobiliser, on injecte 50  $\mu\text{L}$  d'éthanolamine 1 M pH 8,5 (commercialisée par GE Healthcare) afin de bloquer tous les fonctions carboxyles non utilisées pour l'immobilisation des protéines. Enfin on injecte 10  $\mu\text{L}$  d'HCl 10 mM puis 20  $\mu\text{L}$  de NaCl 1 M afin d'enlever toutes les protéines qui ne sont pas couplées de façon covalente.

### 8.2.2.1 Surface d'étude de l'interaction entre la Langérine et la BSA glycosylée avec un trisaccharide composé de $\text{Man}\alpha 1\text{-}3\text{Man}\alpha 1\text{-}6\text{Man}$

Canal de référence : Le canal de référence est fonctionnalisé avec de la BSA (Bovin Serum Albumin). 50  $\mu\text{L}$  sont injectés à une concentration de 60  $\mu\text{g}/\text{mL}$  à un débit de 5  $\mu\text{L}/\text{min}$ . La réponse finale à la fin de l'injection de NaCl permet d'évaluer la quantité de protéine fixée. On obtient un taux de fonctionnalisation aux alentours de 4 600 RU. Une quantité comparable est fixée sur le canal de travail et de référence afin d'effectuer une soustraction des signaux.

Canal de « travail » : Celui-ci est fonctionnalisé avec de la BSA glycosylée en 12 sites avec un trisaccharide composé de  $\text{Man}\alpha 1\text{-}3\text{Man}\alpha 1\text{-}6\text{Man}$  (BSA-Man) lié à la BSA avec 14 atomes d'espacement (Dextra laboratories). La protéine glycosylée est injectée à 60  $\mu\text{g}/\text{mL}$  pendant 5 min (exactement 26  $\mu\text{L}$ ) afin d'obtenir une réponse finale de 4 800 RU.

### 8.2.2.2 Surface d'étude de l'interaction entre la Langérine ECD et gp120<sub>YU2</sub> (expériences complémentaires, chap. 12)

Pour cette étude, la piste référence est fonctionnalisée avec de la BSA comme précédemment décrit, mais en injectant 5  $\mu\text{L}$  à 10  $\mu\text{g}/\text{mL}$ . La réponse finale d'immobilisation est d'environ 700 RU. L'objectif est de masquer la surface par une protéine afin de ne pas excessivement surévaluer les interactions non spécifiques.

Sur la piste de travail, nous injectons 8  $\mu\text{L}$  de gp120<sub>YU2</sub> à 10  $\mu\text{g}/\text{mL}$ . La réponse finale de greffe est de 800 RU.

### 8.2.2.3 Surface d'étude de l'interaction entre la Langérine et l'héparine (expériences complémentaires, cf chap. 12)

Pour l'étude de l'interaction entre la Langérine et l'héparine, nous avons utilisé les outils développés par l'équipe SAGAG de l'IBS (« Structure and Activity of Glycosaminoglycans ») au sein d'une collaboration. Ils ont à leur disposition de l'héparine 6 et 15 kDa biotinylée qui permet une fixation forte, orientée et spécifique sur une surface préalablement immobilisée de streptavidine.

Donc sur la piste référence et la (les) piste(s) de « travail » nous injectons, en utilisant la procédure classique de fonctionnalisation, 50  $\mu$ L de streptavidine à 100  $\mu$ g/mL pour obtenir à la fin de la procédure la saturation de la surface avec la streptavidine qui correspond à une réponse finale de l'ordre de 4 000 RU.

Après la fonctionnalisation des surfaces avec la streptavidine, nous pouvons donc immobiliser l'héparine biotinylée (6 kDa) sur la piste de « travail ». Pour cela, nous injectons, à un débit de 5  $\mu$ L/min et avec le tampon HBS-P comme tampon de course, entre 20 et 65  $\mu$ L (suivant la réactivité de la surface) d'héparine biotinylée à 5  $\mu$ g/mL dans un tampon NaCl 0,3 M par pallier de 10  $\mu$ L jusqu'à obtenir une réponse entre 20 et 30 RU. Afin d'éliminer la fixation non spécifique, nous injectons 20  $\mu$ L de NaCl à 2 M.

### 8.2.2.4 Surface d'étude de l'interaction entre la Langérine ECD et la Langérine S-ECD (expériences complémentaires, chap. 12)

Pour cette étude, nous utilisons le même principe que pour l'immobilisation de l'héparine mais cette fois-ci en utilisant la spécificité d'interaction entre la streptactine et le « Strep Tag II » porté par la protéine.

Donc sur le canal référence et le (les) canal(aux) de « travail » nous injectons, en utilisant la procédure classique de fonctionnalisation, 50  $\mu$ L de streptactine à 100  $\mu$ g/mL pour obtenir à la fin de la procédure la saturation de la surface avec la streptactine qui correspond à une réponse finale de l'ordre de 3 000 RU.

Contrairement à l'héparine, l'affinité de la streptactine pour le « Strep Tag II » n'est pas assez forte pour être stable au cours du temps. Pour résoudre ce problème, nous activons la surface de « travail » avec l'EDC/NHS (50  $\mu$ L) pour coupler covalamment la streptactine et la Langérine S-ECD. Nous immobilisons la Langérine S-ECD en injectant 50  $\mu$ L de protéine à 60  $\mu$ g/mL dans un tampon NaCl 0,3 M pour limiter les interactions non spécifiques. Puis on injecte 50  $\mu$ L d'éthanolamine et 20  $\mu$ L de NaCl 2 M. La réponse d'immobilisation est d'environ 1 000 RU.

Sur ces surfaces, une seule étude a été réalisée afin d'observer si la Langérine est capable d'interagir avec elle-même mais aussi d'observer le pontage de Langérine par l'héparine.

Pour cela, une première injection de 50  $\mu$ L de Langérine ECD à 500 nM a été réalisée

pour observer l'interaction Langérine/Langérine. Après régénération des surfaces avec 20  $\mu\text{L}$  de NaCl 1 M, nous avons injecté, en utilisant la fonction « coinject » qui permet d'injecter successivement deux molécules, 100  $\mu\text{L}$  d'héparine 15 kDa à 500 nM puis 50  $\mu\text{L}$  de Langérine ECD à 500 nM. La régénération s'effectue aussi avec 20  $\mu\text{L}$  de NaCl 1 M. Ensuite l'expérience inverse de la précédente est réalisée, nous avons donc injecté d'abord 50  $\mu\text{L}$  de Langérine puis 100  $\mu\text{L}$  d'héparine 15 kDa.

### 8.2.3 Expériences d'inhibition

#### 8.2.3.1 Etude de l'inhibition de l'interaction entre la Langérine et l'héparine par des fragments d'héparine de différentes tailles et par l'héparine et l'héparane sulfate désulfaté en C6

Le tampon de course utilisé pour les expériences est composé de Tris-HCl 25 mM pH 8, NaCl 150 mM,  $\text{CaCl}_2$  4 mM et de surfactant P20 0,005%. Il est utilisé à un débit de 10  $\mu\text{L}/\text{min}$ .

La concentration en Langérine ECD est fixée à 100 nM et celle des inhibiteurs est de 1  $\mu\text{M}$ . Le volume injecté est de 50  $\mu\text{L}$ . Les inhibiteurs utilisés sont des fragments d'héparine obtenus par digestion enzymatique avec l'héparinase III et purifiés par chromatographie. Ces étapes ont été réalisées par Romain Vivès et Emilie Vassal de l'équipe SAGAG. Les fragments vont d'un hexasaccharide (dp6) à un dp18.

L'héparine et l'héparane sulfate désulfatés en C6 des glucosamines sont obtenus par digestion enzymatique avec HSulf2 puis purifiés par chromatographie. L'analyse et la quantification de désulfatation ont été faites par Romain Vivès et Emilie Vassal.

#### 8.2.3.2 Etude de l'inhibition de l'interaction entre la Langérine et gp120 ou la BSA-Mann par l'héparine

Le tampon de course utilisé pour les expériences est composé de Tris-HCl 25 mM pH 8, NaCl 150 mM,  $\text{CaCl}_2$  4 mM et de surfactant P20 0,005%. Il est utilisé à un débit de 20  $\mu\text{L}/\text{min}$ .

La concentration en Langérine ECD est fixée à 500 nM et celle de l'héparine 15 kDa varie de 7,8 nM à 8  $\mu\text{M}$  pour l'expérience sur la surface BSA-Mann et de 7,8 nM à 2  $\mu\text{M}$  pour l'expérience sur la surface gp120. Le volume injecté de chaque échantillon est de 250  $\mu\text{L}$  et la régénération est effectuée avec 50  $\mu\text{L}$  d'EDTA 50 mM.

### 8.2.3.3 Etude de l'inhibition de l'interaction entre la Langérine et la BSA-Mann par les mono et trisaccharides sulfatés

Le tampon de course utilisé pour les expériences est composé de Tris-HCl 25 mM pH 8, NaCl 150 mM, CaCl<sub>2</sub> 4 mM et de surfactant P20 0,005%. Il est utilisé à un débit de 20 µL/min.

Avant d'effectuer les expériences d'inhibition, il faut s'assurer que les protéines interagissent avec le ligand fixé sur la surface. Pour cela, on injecte à plusieurs reprises, afin d'évaluer la reproductibilité de l'interaction, 100 µL de Lg-ECD à 100 nM. Entre chaque injection, il est impératif de régénérer la surface pour dissocier les complexes formés. Pour cela nous utilisons la capacité de ces lectines (DC-SIGN et la Langérine) à interagir de façon dépendante du calcium avec le mannose. Cette particularité permet de régénérer doucement et entièrement les surfaces par l'injection de 20 µL d'EDTA 50 mM.

Pour évaluer l'affinité des différents sucres à tester, nous estimons la concentration pour laquelle un sucre donné inhibe à 50% l'interaction de la protéine avec les trimannoses exposés à la surface de la BSA (IC<sub>50</sub>).

Pour effectuer ce type d'estimation, il faut injecter un mélange de protéines-sucres, préalablement incubés, composé d'une concentration fixe en protéine (100 nM) et de concentration variable en sucre (concentration en sucre commençant de 0 M). On peut donc, à chaque fin d'injection du mélange, relever la réponse occasionnée par l'interaction des protéines non inhibées et ainsi estimer, par rapport à une solution ne contenant que la protéine, le pourcentage d'inhibition pour chaque concentration en sucre. Les expériences ont été réalisées en duplicat.

Les concentrations maximales en sucres utilisées sont différentes pour chacun. La concentration maximale dépend de la quantité de sucre disponible. Les gammes en sucre utilisées sont représentées dans le tableau 8.1 :

Saccharide	Gamme utilisée (mM)
GlcNac	0,047-24
GlcNac6S	0,023-6
GlcNS6S	0,023-6
Trila	0,003-0,75

Tableau 8.1 – Gamme utilisée des différents sucres pour les expériences d'inhibitions.

### 8.2.4 Microscopie électronique à coloration négative (expériences complémentaires, chap. 12)

Les expériences de microscopie électronique en transmission ont été réalisées par Guy Schoehn à l'IBS. Le microscope utilisé est un Philips CM12.

La méthode utilisée pour la visualisation est la coloration négative. Elle consiste à adsorber les protéines sur une grille métallique recouverte d'un film de carbone fin. Puis d'enlever l'excès de solution aqueuse par un buvard. Ensuite, une solution contrastante, dans notre cas l'acétate d'uranyle, est ajoutée sur la grille pendant quelques secondes puis absorbée avec un buvard. L'acétate d'uranyle va se fixer autour des molécules adsorbées. Par sa forte masse, le contrastant dévie les électrons et ainsi la protéine apparaîtra plus claire que ce qui l'entoure, d'où le nom de coloration négative.

Deux types d'échantillons ont été observés :

- La Langérine ECD seule à une concentration de 0,1  $\mu\text{g/mL}$ .
- La Langérine ECD (0,1  $\mu\text{g/mL}$ ) en présence d'héparine (1  $\mu\text{g/mL}$ ).



# Chapitre 9

## Cristallographie

La cristallographie est la science qui se consacre à l'étude des substances cristallines à l'échelle atomique. Les propriétés physico-chimiques d'un cristal sont étroitement liées à l'arrangement spatial des atomes dans la matière. L'état cristallin est défini par un caractère périodique et ordonné à l'échelle atomique ou moléculaire. Le cristal est obtenu par translation dans toutes les directions d'une unité de base appelée maille élémentaire.

### 9.1 Le synchrotron

Le synchrotron est un instrument permettant l'accélération à haute énergie de particules stables chargées (Figure 9.1). Il contraint les particules à tourner en rond entre les pôles d'électro-aimants disposés en anneau. Les particules sont accélérées à chaque tour. Pour les maintenir confinées dans le cercle, on ajuste le champ magnétique à l'énergie atteinte par les particules. En raison de la faible masse des électrons, la décélération occasionnée par la courbure de leur trajectoire génère une onde électromagnétique, le rayonnement synchrotron. Ce rayonnement est collecté à différents endroits du tore, les lignes de lumière. Chaque faisceau lumineux rencontre ensuite des lentilles, miroirs ou monochromateurs afin de sélectionner la gamme de longueurs d'onde d'intérêt et de modifier les caractéristiques du faisceau (taille, divergence) qui sera utilisé dans l'expérience. « Au bout » de chaque ligne de lumière est monté un échantillon de matière servant de cible. Les électrons éjectés lors de l'interaction du faisceau incident avec la cible sont détectés par des appareils de mesure ponctuels, linéaires ou bidimensionnels (caméra CCD, image plate). Suivant la taille de l'anneau, jusqu'à des dizaines d'expériences peuvent être menées simultanément.

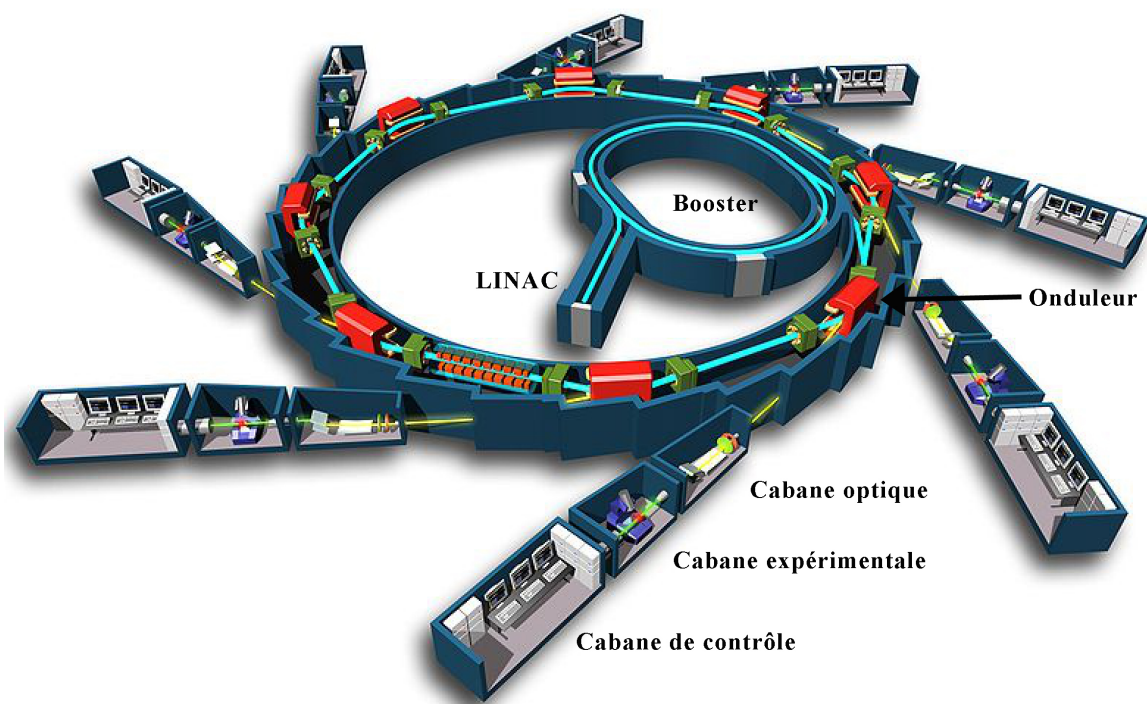


Figure 9.1 – Schéma représentant les équipements de bases du synchrotron SOLEIL (Paris). Source : Synchrotron SOLEIL

## 9.2 La cristallogénèse

Afin d'obtenir des cristaux de protéines, il est nécessaire de réduire la solubilité de celles-ci. Pour cela, la technique la plus répandue est la cristallisation par diffusion en phase vapeur. Cette technique consiste à créer une différence de concentration entre une solution protéique et une solution précipitante dans un milieu hermétiquement clos, et de tendre vers un retour à l'équilibre par diffusion de vapeur de l'eau.

Différents procédés sont utilisés, comme la méthode de goutte assise ou suspendue. C'est cette dernière qui est utilisée dans ce travail de thèse.

La solution de protéine est diluée par mélange avec un volume équivalent de la solution du réservoir pour former une goutte pendant vers le bas, disposée sur la face inférieure d'une lamelle située au-dessus de la solution « réservoir », laquelle contient une concentration très élevée de précipitant. Le retour à l'équilibre des concentrations est atteint lentement par diffusion de vapeur de l'eau, de la goutte vers le réservoir. La diminution du volume de la goutte entraîne en même temps l'augmentation des concentrations en agents précipitants et en protéine. La forte différence de volume entre la goutte et le réservoir sert à accélérer le retour à l'équilibre (Figure 9.2).

Le volume initial de la goutte est de 2  $\mu\text{L}$  constitué de 1  $\mu\text{L}$  de la solution de précipitant

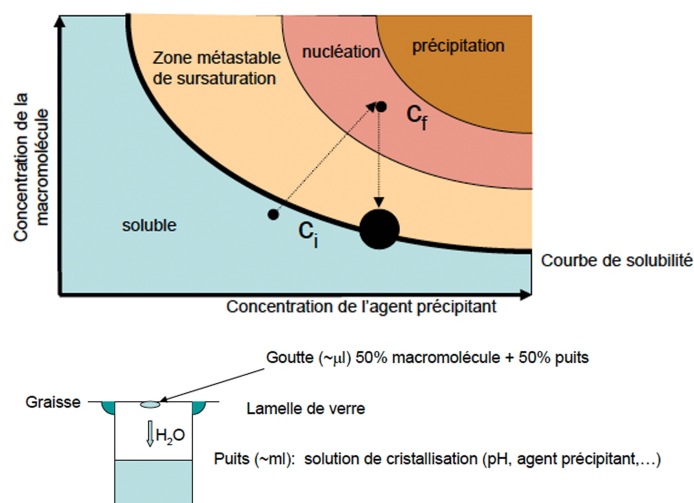


Figure 9.2 – Schéma représentant le diagramme de phase (haut) pour la méthode de goutte suspendue (bas). D’après le livre « Cristallographie » réalisé par Dieter Schwarzenbach et Gervais Chapuis.

et de 1  $\mu$ L de la solution de protéine, tandis que le volume de la solution de précipitant équivaut à 1 ml.

Au fur et à mesure l’eau contenue dans la goutte va diffuser sous forme de vapeur dans la solution réservoir, ainsi l’équilibration s’effectue par évaporation des composés dit « volatiles », comme l’eau jusqu’à ce que les tensions de vapeur à la surface des deux liquides soient identiques.

La concentration de la protéine dans la goutte va alors augmenter, du fait de la diminution du volume de la goutte, ce qui va entraîner la nucléation (Figure 9.2). Pour la croissance des cristaux, la protéine cristallisera dans une durée de temps variable selon la protéine et les conditions utilisées.

### 9.3 Obtention des cristaux de Langérine CRD

La cristallisation a été réalisée avec une solution protéique concentrée à 15 mg/mL de Lg-EC dans son tampon de purification et avec une solution de précipitant contenant Hepes 100 mM pH 7 ;  $MgCl_2$  100 mM ; NaCl 150 mM et de 18% à 20% de PEG 3350. La croissance des cristaux prend environ deux mois, c’est le temps nécessaire à la protéolyse de la protéine (elle est aussi observée dans le cas de la cristallisation du CRD, voir chap.13). Des cristaux ont aussi été obtenus dans les mêmes conditions de cristallisation (en tampon Hepes 100 mM pH 7 ou MES 100 mM pH 6) mais en présence de 70 mM de dipicolinate d’euporium. Dans ces conditions le volume de goutte initial était de 3  $\mu$ L.

## 9.4 Le trempage avec les sucres

Les cristaux servant au trempage avec les sucres sont préalablement nettoyés dans la solution du puit et sont ensuite incubés dans une goutte de trempage contenant le sucre. Cette goutte de trempage est préparée par le mélange de 1  $\mu\text{L}$  d'une solution contenant les mêmes concentrations en tampon et en sel que la solution de cristallisation mais avec une concentration en PEG 3350 de 35% ce qui permet la cryoprotection directe des cristaux et de 1  $\mu\text{L}$  de la solution de sucre. La concentration finale en sucre est d'environ 10 mM pour les trois sucres dérivant du GlcNac et de 5 mM pour le Tri1a. Ces gouttes sont équilibrées au dessus d'un puit contenant 1 mL de la solution cryoprotectante pendant 4 h. Ensuite les cristaux (3 cristaux pour chaque sucre) sont placés dans les gouttes pendant 1 h. Enfin les cristaux sont indépendamment placés dans une boucle et congelés à l'azote liquide.

## 9.5 Co-cristallisation

La concentration de saccharide dans les gouttes du trempage est d'environ 10 mM. Les cristaux sont donc incubés dans la condition de cristallisation complétée des différents saccharides et d'une concentration en PEG supérieure permettant par la suite la cryoprotection. Les cristaux sont congelés à l'azote liquide après leur rapide rinçage dans la solution cryoprotectante déplétée de sucre. La diffraction a été réalisée sur BM30a FIP de l'ESRF de Grenoble.

Saccharides	DUA-GlcNac	DUA-GlcNac6S	DUA-GlcNS6S	DUA2S-GlcNS6S	DUA-GalNac	DUA-GalNac6S
Nombre de conditions testées	540 (6*90)	540	540	540	540	540
Pistes obtenues	0	0	0	0	0	1
Composition des conditions positives						0,1 M ADA pH6,5; 0,1 M sulfate de lithium; 1 M sulfate de magnésium
Reproductibilité en $\mu\text{goutte}$						Non
Composition des conditions positives ( $\mu\text{goutte}$ )						0,1 M ADA pH7; 0,1 M sulfate de lithium; 1,4 M sulfate de magnésium

Tableau 9.1 – Résumé de la cristallogénèse de la Langérine CRD avec les différents disaccharides.

## 9.6 La cristallographie en biologie structurale

Une fois des cristaux de taille suffisante obtenus (quelques dizaines à quelques centaines de microns), on les place dans un faisceau de rayons X monochromatique, produit par une

source adaptée, un générateur de laboratoire ou une source de rayonnement synchrotron, ce dernier produisant des faisceaux plus intenses. Les rayons X qui frappent le cristal sont diffusés par les atomes le constituant. Les rayons X diffusés interfèrent entre eux. Dans un cristal où les molécules sont ordonnées, les interférences vont être constructives dans certaines directions, qui forment le phénomène de diffraction, et destructives dans d'autres. Avec un détecteur, on enregistre des images du motif de diffraction (interférences constructives) obtenu qui est constitué de taches régulièrement espacées, pour différentes orientations du cristal (Figure 9.3). L'espacement des taches renseigne sur la maille du cristal, c'est-à-dire les dimensions de l'élément de volume qui est régulièrement répété dans l'espace pour donner naissance au cristal, mais aussi sur les éléments de symétrie du cristal.

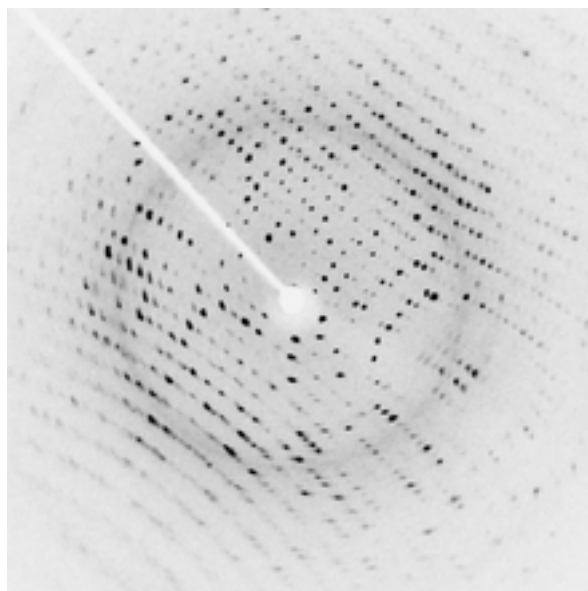


Figure 9.3 – Exemple de motif de diffraction pour le lysosyme. Cliché réalisé par Romain Talon (IBS).

Les intensités des taches contiennent quant à elles toute l'information sur la densité électronique en chaque point de l'espace des macromolécules présentes dans la maille. La phase du signal associé à chaque tache est cependant perdue et il faut alors la reconstituer. En effet les intensités des taches correspondent au carré du module du facteur de structure d'indice de Miller  $hkl$  (équation 1 :  $I_{hkl} = |F(hkl)|^2$ ). Or  $F(hkl) = |F(hkl)|e^{-i\varphi(hkl)}$  (équation 2), donc à partir des intensités des taches nous obtenons le module du facteur de structure ( $|F_{hkl}|$ ) mais nous n'avons pas l'information de phase ( $\varphi(hkl)$ ). Or il faut effectuer la transformée de Fourier du facteur de structure pour obtenir une carte de densité électronique (équation 3 :  $\rho(x, y, z) = \frac{1}{V} \sum_{hkl} F(hkl)$ ). Ce qui est impossible si nous n'avons pas l'information de phase. Pour récupérer l'information de phase, plusieurs

techniques sont possibles :

- Reconstruction des phases par des méthodes *ab initio* : ces méthodes consistent à générer des différences dans un cristal ou plusieurs cristaux de la même protéine par l'introduction d'une sous-structure. Pour cela, on introduit généralement des atomes lourds ou on remplace les méthionines et les cystéines par des sélénométhionines et sélénocystéines. On utilise leurs propriétés physiques. Dans le cas d'atomes lourds, on utilise le fait qu'ils sont riches en électrons, et pour les acides aminés séléniés, on utilise le signal anomal. Ceci permet de retrouver les positions de ces atomes dans le cristal et ainsi obtenir l'information de phase manquante. Nous n'avons pas utilisé ces méthodes pour déterminer les différentes structures de ce travail mais la méthode suivante.
- Ou par la méthode dite de remplacement moléculaire : Le remplacement moléculaire est une méthode utilisant la structure d'une protéine homologue lorsque celle-ci est connue. On essaye de placer la structure homologue dans la maille du cristal de la protéine inconnue et on calcule la diffraction théorique que l'on compare à celle qui est observée. Ceci permet de calculer des phases initiales pour la protéine de structure inconnue.

Une fois les phases reconstruites, on peut remonter à la densité électronique par transformation de Fourier des intensités de diffraction. On peut alors reconstruire la molécule dans la densité électronique en utilisant des programmes automatiques ou manuellement à l'aide de logiciels graphiques interactifs, ce qui permet d'obtenir les coordonnées cartésiennes de chaque atome de la protéine inconnue. A partir du modèle construit, on peut calculer des facteurs de structure (F) que l'on compare aux facteurs de structure mesuré expérimentalement, ce qui permet d'améliorer progressivement le modèle, une étape appelée « affinement ».

(source : inspiré du livre « Cristallographie » réalisé par Dieter Schwarzenbach et Gervais Chapuis)

## 9.7 L'enregistrement des données

Tous les enregistrements décrits dans cette section ont été réalisés à l'ESRF (European Synchrotron Radiation Facility, Grenoble) sur la ligne de lumière BM30a (FIP) possédant un détecteur ADSC Q315r.

GlcNac :

- Longueur d'onde utilisée : 0,979 Å
- Nombre d'images et pas angulaire entre chaque : 200 images avec un pas angulaire de 0,5 degré.

- Temps d'exposition : 25 sec/image
- Distance du détecteur : 220 mm

GlcNac6S :

- Longueur d'onde utilisée : 0,979 Å
- Nombre d'images et pas angulaire entre chaque : 200 images avec un pas angulaire de 0,5 degré.
- Temps d'exposition : 25 sec/image
- Distance du détecteur : 201,8 mm

GlcNS6S :

- Longueur d'onde utilisée : 0,979 Å
- Nombre d'images et pas angulaire entre chaque : 200 images avec un pas angulaire de 0,5 degré.
- Temps d'exposition : 15 sec/image
- Distance du détecteur : 220,7 mm

Trila :

- Longueur d'onde utilisée : 0,979 Å
- Nombre d'images et pas angulaire entre chaque : 200 images avec un pas angulaire de 0,5 degré.
- Temps d'exposition : 20 sec/image
- Distance du détecteur : 220,7 mm

## 9.8 Protocole utilisé pour la détermination des structures

Les données de diffraction ont été intégrées avec le programme XDS [58]. Après vérification, le groupe d'espace est le même que pour la Langérine S-CRD (protéine de départ Lg-EC), c'est à dire  $P4_2$ . Pour chaque jeu de données, on utilise un jeu de référence unique qui permet d'éliminer tout problème d'indexation due au groupe d'espace. Ce jeu de référence est le jeu de données traité par Michel Thépaut qui a conduit à la détermination de la structure de la Langérine S-CRD [100]. A la fin du traitement par XDS, celui-ci crée un fichier XDS\_ASCII.HKL qui contient toutes les intensités des réflexions enregistrées pour chaque image et permet de générer le premier fichier mtz. Pour cela, nous utilisons un script créé par Richard Kahn qui permet de combiner la création du mtz par le programme combat, la mise à l'échelle locale par le programme Scala et la transformation des intensités en module du facteur de structure avec le programme Truncate.

Les structures ont été déterminées par la méthode de remplacement moléculaire. Nous avons ainsi créé un modèle initial à partir de la structure de la Langérine CRD exis-

tante (code PDB : 3C22). Cette structure a été modifiée en supprimant les molécules d'eau, d'ions et les doubles conformations des chaînes latérales. Ensuite nous appliquons à chacun des résidus un taux d'occupation de 1 et un facteur d'agitation thermique B correspondant à la valeur de Wilson obtenue à l'étape Truncate. Enfin, nous créons, à partir de cette structure modifiée, une nouvelle structure faite exclusivement d'alanine (structure poly-alanine). L'objectif est d'éliminer au maximum le biais provenant du modèle dans l'information de phase, en particulier celles des chaînes latérales.

Le programme Phaser [70] (de la suite ccp4 [16]) est utilisé pour faire le remplacement moléculaire. Ensuite, nous avons utilisé le programme DM [16] à partir du fichier mtz généré par Phaser afin d'améliorer l'information de phase par aplatissement de solvant. Pour cela, une valeur de la quantité de solvant doit être implémentée, nous avons donc pris l'estimation donnée dans le fichier truncate.log en l'abaissant de 5% pour éviter une perte d'information. Enfin, à partir des phases améliorées on effectue une reconstruction automatique du modèle par le programme Buccaneer en utilisant le fichier mtz généré par DM et la séquence de la Langérine.

Afin d'éviter tout biais dans la comparaison des différentes structures au niveau de l'affinement et par la suite, nous utilisons le même jeu de réflexions pour évaluer le facteur Rfree. Pour cela nous avons utilisé le programme CAD pour transférer le jeu de réflexions utilisés, pour évaluer le Rfree, lors de l'affinement de la structure de la Langérine CRD (code PDB : 3C22).

Nous avons utilisé le programme Phenix.refine pour l'affinement de la structure. La première étape consiste à affiner la position de chaque atome (x, y, z), ainsi que le facteur d'agitation thermique B. Nous réalisons aussi un recuit simulé qui permet d'optimiser l'énergie du système. Ensuite et comme après chaque étape de 5 cycles d'affinement, une observation de la structure affinée dans la nouvelle carte de densité est réalisée durant laquelle on effectue les modifications manuelles pour placer correctement les résidus dans la densité électronique. Pendant ces étapes et si nécessaire, on ajoute des doubles conformations pour les chaînes latérales des résidues mobiles. On observe aussi des zones de densité non construite correspondant au calcium du site de liaison ainsi que des sucres en liaison avec celui-ci. Après ces étapes, on effectue plusieurs cycles d'affinement en optimisant les positions des atomes, les facteurs B individuels, le taux d'occupation et on implémente aussi l'ajout de molécules d'eau. Après ces étapes, on vérifie grâce à la carte de densité  $F_O - F_C$  (facteur de structure observé - facteur de structure calculé) la position des sucres et de son taux d'occupation. On vérifie aussi le placement des molécules d'eau, c'est à dire que les molécules d'eau dans une densité de moins de 1 sigma sont enlevées ainsi que celles qui sont trop proches de résidues ( liaisons hydrogènes entre 2,5 Å et + 3,5 Å). On vérifie aussi par le graphique de Ramachandran la bonne géométrie de tous



les résidus. A la fin de ces étapes, on ajoute l'anisotropie sur les atomes de calcium et on place d'autres ions (les molécules d'eau placées dans certaines densité ne procure pas assez de densité électronique, par l'observation de l'environnement on peut discriminer si les atomes à placer sont des ions magnésium ou chlore). En dernière étape (réalisée que pour le GlcNac6S) on effectue une étude TLS qui permet un affinement des facteurs B pour des groupes de résidus.

---

## Quatrième partie

### Résultats

---

# Chapitre 10

## Obtention des outils protéiques

### 10.1 Expression et purification des Lg-S-CRDs

Les plasmides, possédant un gène de résistance à la kanamycine et contenant le gène des Lg-S-CRDs sauvage et mutantes précédé des séquences codantes pour le « Strep Tag II » et le peptide OmpA, sont transformés dans des bactéries *E. coli*. Les bactéries correctement transformées sont sélectionnées sur boîtes de Petri composées de milieu LB supplémenté de kanamycine.

L'expression des Lg-S-CRD est réalisée de façon recombinante dans des bactérie *E. coli* de souche BL21(DE3). La particularité de cette souche de bactérie est l'absence de certaines protéases. L'ajout en N-terminal de la séquence bactérienne du peptide signal OmpA permet l'adressage de la synthèse protéique à la membrane et l'export des protéines néo-synthétisées dans le périplasme. Après le transfert des néo-protéines dans le périplasme, le peptide est clivé. Cet adressage dans ce compartiment oxydant permet la bonne structuration des protéines par l'établissement des ponts disulfures essentiels à la structure du domaine CRD (voir introduction sur les lectines). Ceci permet aussi d'obtenir après la lyse des culots bactériens une protéine soluble pouvant être directement purifiée sans passer par des étapes fastidieuses de renaturation, indispensables pour DC-SIGN par exemple.

L'expression des protéines est réalisée à 20°C pour limiter la formation de corps d'inclusion. En effet, les Lg-S-CRDs acheminées dans le périplasme peuvent acquérir une bonne structuration. Mais si la production des protéines est trop soutenue, le système d'export s'engorge et il y a formation de corps d'inclusion. L'expression à 20°C ralentit la synthèse protéique et améliore le rendement d'export périplasmique où le repliement se fera correctement. Une vérification de l'expression des protéines est réalisée par gel SDS-PAGE à 15% (Figure 10.1). Sur le gel, on observe bien une bande d'intensité forte apparaissant aux alentours de 15 kDa correspondant à la Lg-S-CRD.

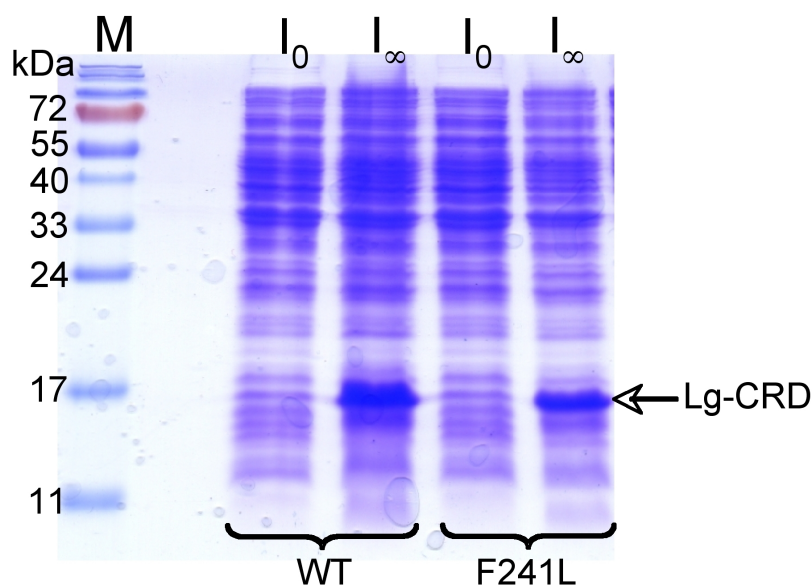


Figure 10.1 – SDS-PAGE 15% montrant l'expression de Lg-S-CRD. M : marqueur de poids moléculaire.  $I_0$  : Extrait de bactéries avant induction.  $I_\infty$  : Extrait de bactéries à la fin de l'expression de Lg-S-CRD.

A la fin de l'expression, les bactéries sont centrifugées puis resuspendues pour effectuer la lyse de celles-ci par sonication. Après cette étape, la solution est ultracentrifugée pour éliminer les membranes et autres agrégats. Le surnageant récupéré est purifié par chromatographie d'affinité sur une colonne de streptactine. En effet, les Lg-S-CRD possèdent en N-ter une étiquette de 8 acides amines dénommée « Strep Tag II » qui possède une forte affinité pour le site de fixation de la biotine de la protéine dérivée de la streptavidine (la streptactine). La colonne est tout d'abord équilibrée dans le tampon A (Tris-HCl 25 mM pH8 ; NaCl 150 mM ;  $\text{CaCl}_2$  4 mM), puis le surnageant de la casse des bactéries est injecté et la colonne est lavée par plusieurs volumes de tampon A. A la fin de cette étape, toutes les protéines ne possédant pas le « Strep Tag II » passent au travers de la colonne. Les Lg-S-CRD sont éluées par un tampon contenant un compétiteur du « Strep Tag II » de plus forte affinité, la Desthiobiotine (un analogue de la biotine de moindre affinité permettant ultérieurement la réutilisation de la colonne). Sur le chromatogramme, on observe donc une première phase pendant laquelle beaucoup de protéines passent au travers de la colonne et une phase d'éluion qui correspond donc aux protéines ayant une affinité spécifique pour la colonne (Figure 10.2).

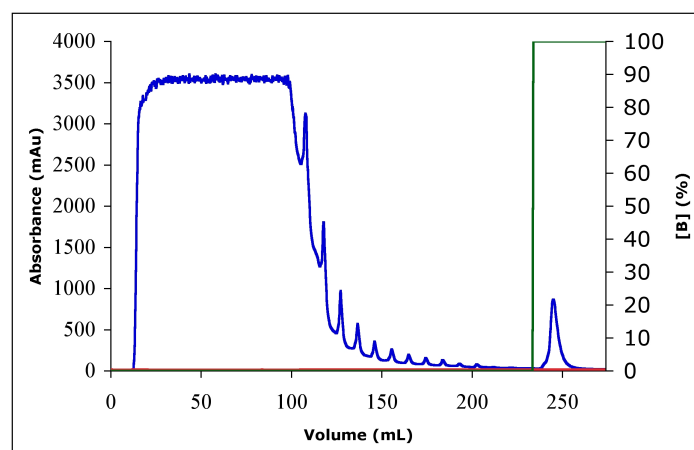


Figure 10.2 – Chromatogramme de purification de Lg-S-CRD. Bleu : Absorbance à 280 nm. Vert : Concentration en % du tampon d'élution (axe des ordonnées secondaires).

Les fractions éluées sont récupérées et les protéines sont concentrées afin de réaliser une deuxième étape de purification en utilisant la capacité de la Lg-S-CRD à reconnaître le mannose. L'affinité du domaine CRD n'est pas suffisant pour que les protéines soient fixées sur la colonne, on sélectionne ainsi, par le suivi de l'absorbance à 280 nm, les protéines ayant un retard d'élution. En effet, lorsque l'on observe le chromatogramme de purification sur la colonne de mannose, on visualise un premier pic correspondant aux protéines non fonctionnelles sortant au volume totale de la colonne, puis un large pic correspondant aux protéines retardées et éluées par 2 à 4 volumes de colonne supplémentaire (Figure 10.3 A). A la fin de cette étape de purification, on effectue un SDS-PAGE 15% pour vérifier la pureté des protéines (Figure 10.3 B). Sur ce gel, on observe bien une seule bande à environ 15 kDa correspondant à la Lg-S-CRD.

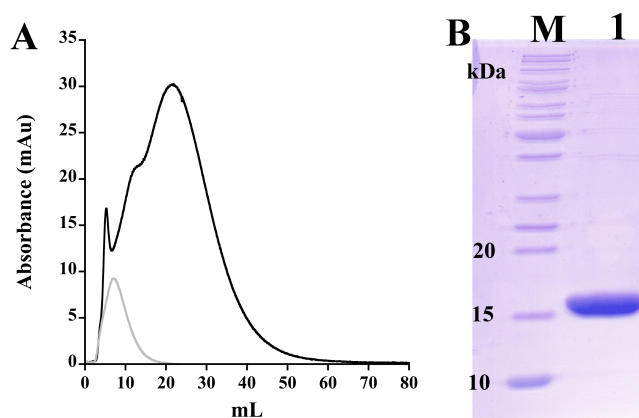


Figure 10.3 – A : Chromatogramme de purification sur colonne de mannose. Noir : échantillon concentré de Lg-S-CRD. Gris : contrôle négatif (BSA). B : SDS-PAGE 15% des protéines purifiées. M : Marqueur de poids moléculaires. 1 : extrait purifié de Lg-S-CRD.

La quantité de Lg-S-CRD par litre de culture initiale peut varier de 0,5 à 2 mg de protéines. Cette variation n'est pas due à l'expression mais plutôt à la quantité de protéine correctement repliée, et donc soluble, dans le périplasme bactérien.

## 10.2 Expression et purification des Lg-ECDs

La construction plasmidique codant pour les protéines Lg-ECs ne possède aucun signal d'adressage au périplasme ni de Tag pour la purification. Des tests d'expression préalables ont montré l'incapacité de produire le domaine ECD soluble dans le périplasme. La formation de corps d'inclusion dans le cytoplasme permet d'obtenir une plus grande quantité de protéines que dans le périplasme.

L'expression des protéines est réalisée chez les bactéries *E. Coli* BL21(DE3) mais cette fois-ci à 37°C (Figure 10.4).

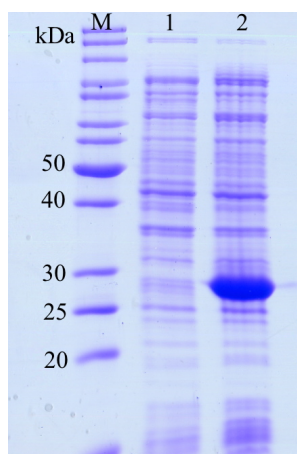


Figure 10.4 – SDS-PAGE 12% de l'expression de Lg-ECD. M : Marqueur de poids moléculaires. 1 : Extrait total avant induction. 2 : Extrait total à la fin de l'induction

Cette expression massive de protéines forme des corps d'inclusion constitués principalement de Lg-ECD non structuré. Ceci permet d'obtenir une grande quantité de protéine mais nécessite impérativement de réaliser des étapes de renaturation pour obtenir la protéine oligomérique et fonctionnelle.

La première étape consiste à récupérer les corps d'inclusion après la lyse des bactéries par une ultracentrifugation. Le surnageant est éliminé car il contient toutes les protéines bactériennes solubles. Le culot, d'une couleur blanchâtre typique de corps d'inclusion, est resuspendu dans un tampon contenant de l'urée 2 M et du triton X100 1% qui permet de solubiliser et d'éliminer des protéines contaminantes agrégées aux corps d'inclusions de Lg-ECD. Cette étape est forcément nécessaire car on élimine énormément de contaminants même si l'on perd un peu de Lg-ECD à cette étape (Figure 10.5, 3). Le culot obtenu après



ultracentrifugation est nettoyé une deuxième fois avec le tampon de lyse des bactéries pour s'assurer qu'aucune protéine soluble ne reste. Ensuite, il est nécessaire de solubiliser les Lg-ECs sans amener les débris de membranes. Pour cela, le culot obtenu une nouvelle fois par ultracentrifugation est resuspendu dans un tampon contenant du chlorure de guanidinium 6 M, NaCl 150 mM, Tris pH 8 25 mM et une faible quantité de  $\beta$ -mercaptoethanol. L'étape la plus importante est la « Flash » dilution. Cette méthode permet de diluer rapidement la guanidine qui entoure et solubilise les protéines pour que ces dernières se structurent dans un état énergétique faible. Pour cela un goutte à goutte est réalisé dans une solution contenant une forte concentration de Tris (210 mM) et de chlorure de sodium (2,5 M). Un effet important sur le rendement de renaturation a été observé par l'utilisation de concentration élevée de Tris (210 mM par rapport à 150 mM). Le calcium quant à lui est présent pour aider la structuration du site de liaison du CRD. Après renaturation, des dialyses permettent de revenir à un échantillon contenant de plus faible concentration en sel et en tampon. L'échantillon final est ultracentrifugé pour éliminer les agrégats résiduels (Figure 10.5). Le surnageant doit être ensuite purifié.

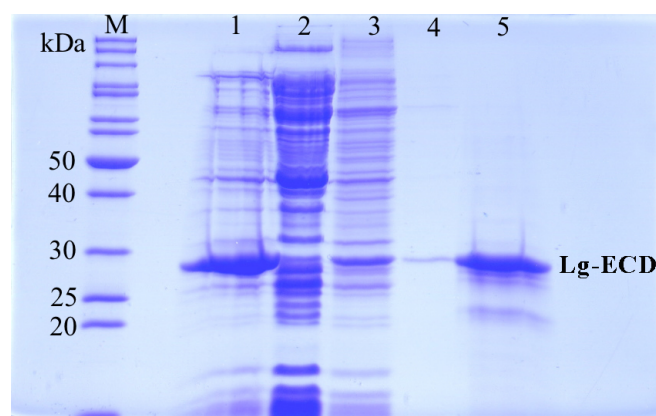


Figure 10.5 – Gel SDS-PAGE 12% des étapes de renaturation des protéines. M : Marqueur de poids moléculaires. 1 : Extrait total. 2 : Surnageant après le premier lavage des corps d'inclusion (tampon : Tris-HCl pH8 25 mM ; NaCl 150 mM). 3 : Surnageant après le deuxième lavage des corps d'inclusion (tampon : Tris-HCl pH8 25 mM ; NaCl 150 mM ; urée 2 M ; triton X100 1%). 4 : Surnageant après le troisième lavage des corps d'inclusion (tampon : Tris-HCl pH8 25 mM ; NaCl 150 mM). 5 : 'Fraction soluble des corps d'inclusion solubilisés (tampon : Tris-HCl pH8 25 mM ; NaCl 150 mM, chlorure de guanidinium 6 M ;  $\beta$ -mercaptoethanol 0,01%).

La purification est réalisée par chromatographie d'affinité sur colonne de Mannan-Agarose (Figure 10.6). Cette stratégie de purification permet de sélectionner les protéines fonctionnelles et oligomériques car, comme pour les Lg-S-CRDs, les protéines non fonctionnelles ne sont pas capables de reconnaître le mannose évidemment et les protéines monomériques ne possèdent pas l'affinité nécessaire pour se fixer sur la colonne. Cette

stratégie permet donc une double réponse à nos problèmes qui sont d'obtenir des protéines fonctionnelles et oligomériques de Lg-ECD, en l'occurrence trimériques. Sur le chromatogramme on observe bien un pic d'élution lorsque du tampon supplémenté en EDTA (et non en calcium) est passé sur la colonne. Ceci montre bien le type de reconnaissance dépendante de l'atome de calcium car l'EDTA est un fort chélatant des métaux divalents.

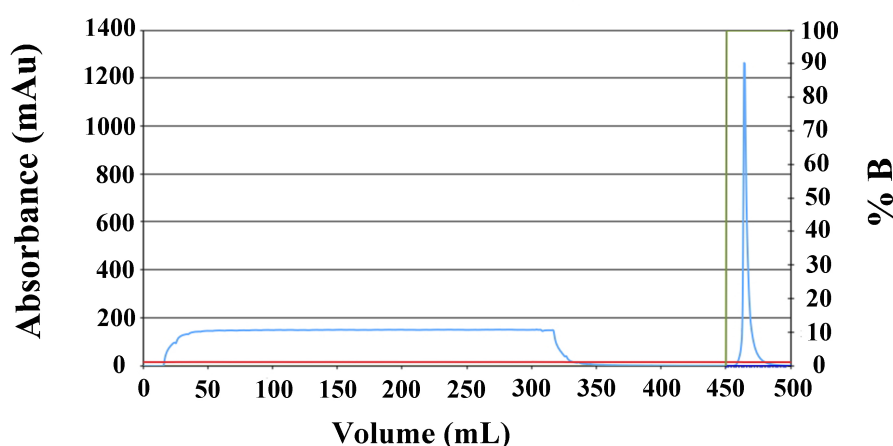


Figure 10.6 – Chromatogramme de purification de Lg-ECD sur une colonne de Mannan-Agarose. Vert : pourcentage de tampon B. Rouge : conductivité.

Dans une étude précédente réalisée au laboratoire, il fût montré qu'après cette étape on récupère une seule espèce de Lg-EC qui est trimérique. Ce résultat a pu être obtenu en analysant l'échantillon purifié et concentré par ultracentrifugation analytique (Figure 10.7)[100]. Cette expérience montre pour deux concentrations (1 mg/mL et 0,1 mg/mL) que l'homogénéité de l'échantillon est parfaite et que le trimère est stable malgré la dilution.

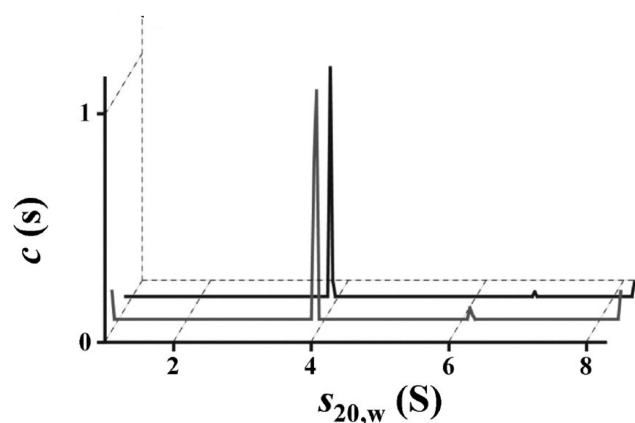


Figure 10.7 – Donnée de sédimentation obtenue par ultracentrifugation analytique de Lg-ECD [100]. Deux concentrations de Lg-ECD ont été analysées, 1 mg/ml (pic du premier plan) et 0,1 mg/mL (pic du second plan).  $s_{20,w}$  : coefficient de sédimentation.  $C(s)$  : concentration normalisée de l'espèce en solution.

Le rendement de renaturation est de 10% environ ce qui est très acceptable sachant que pour 1 L de culture de bactérie on obtient environ 100 mg de protéines en corps d'inclusion (calculé avec l'échantillon solubilisé en chlorure de guanidinium).

---

# Chapitre 11

## Etude de l'effet des mutations F241L et W264R sur la Langérine et le granule de Birbeck

Des études menées par Jenny Valladeau et Pauline Verdick ont montré que des mutations de la Langérine pouvait dramatiquement affecter la formation et la structure des granules de Birbeck (Figure 6.7). De plus, la mutation W264R étant une forme polymorphique humaine, l'identification des conséquences sur le fonctionnement de la Langérine peut sembler utile du fait de l'implication de cette protéine vis-à-vis du VIH. Deux simples mutations dans le domaine de reconnaissance des hydrates de carbones conduisent pour l'une à la perte des granules de Birbeck (mutation W264R, Figure 6.7, C), et pour l'autre à une modification radicale de la structure (mutation F244L, Figure 6.7, B). Par la localisation des mutations dans le CRD, et la nécessité de sa présence pour la structure du granule[100], l'étude de ces mutations pourrait apporter des informations sur la formation des BGs. Des études de biologie cellulaire, de biochimie et de biologie structurale ont été menées tout d'abord pour vérifier l'effet des mutations sur la protéine humaine à l'échelle cellulaire pour ensuite se focaliser sur les effets intra-moléculaires de ces mutations sur la Langérine.

Ce travail a permis :

- D'expliquer la conséquence de la mutation W264R qui déstructure le domaine CRD par la déstabilisation d'une poche hydrophobe.
- De valider l'effet de la mutation F244L sur la structure des BGs pour la protéine humaine (F241L).
- D'identifier l'influence de la mutation F241L dans l'organisation de la structure quaternaire de la Langérine ECD.
- De mettre donc en évidence que le coiled-coil n'est pas le seul domaine responsable

---

de la trimérisation de la protéine.

Ce travail a permis de déterminer pour la première fois des valeurs d'affinité de la Langé-rine CRD pour la glycoprotéine d'enveloppe 120 du VIH.

Dans ce travail, l'établissement de collaborations a été nécessaire notamment pour l'étude de biologie cellulaire, réalisée par Jenny Valladeau-Guillemont et Colette Dezutter-Dambuyant du centre de recherche en cancérologie de Lyon, l'étude SAXS, expertisée par Dominique Durand et Patrice Vachette de l'institut de Biochimie et Biophysique Moléculaire et Cellulaire de l'Université Paris XI. Enfin, les compétences en spectrométrie de masse de Julien Marcoux, ancien doctorant au sein de l'équipe, ont été utilisées.

## Alteration of langerin oligomerisation state affects gp120 recognition and Birbeck granules formation: effect of the F241L mutation.

Eric Chabrol<sup>†¶¶</sup>, Michel Thépaut<sup>†¶¶</sup>, Colette Dezutter-Dambuyant<sup>‡</sup>, Corinne Vivès<sup>†¶¶</sup>, Julien Marcoux<sup>†¶¶</sup>, Richard Kahn<sup>†¶¶</sup>, Jenny Valladeau-Guilemond<sup>‡</sup>, Patrice Vachette<sup>¥§</sup>, Dominique Durand<sup>¥§</sup>, and Franck Fieschi<sup>†¶¶#\*</sup>.

<sup>†</sup>Institut de Biologie Structurale, Université Joseph Fourier-Grenoble, 41 rue Jules Horowitz, 38027 Grenoble, France;

<sup>¶</sup>CEA, DSV, 38027 Grenoble France;

<sup>¶¶</sup>CNRS, UMR 5075, 38027 Grenoble, France;

<sup>#</sup>Institut Universitaire de France, 103 boulevard Saint-Michel, 75005 Paris, France;

<sup>‡</sup>Centre Léon Bérard-UMR INSERM 1052-CNRS 5286, Centre de recherche en Cancérologie de Lyon, 69008 Lyon, France;

<sup>¥</sup>Institut de Biochimie et Biophysique Moléculaire et Cellulaire, Université Paris XI, Bâtiment 430, 91405 Orsay, France;

<sup>§</sup>CNRS, UMR 8619, 91405 Orsay, France.

\* To whom correspondence should be addressed: Franck Fieschi, Institut de Biologie Structurale, 41 rue Jules Horowitz, F-38027 Grenoble Cedex 1, Tel.: (33) 438789177; Fax: (33) 438785494; Email: Franck.Fieschi@ibs.fr

---

### Introduction

Dendritic cells (DCs) are professional antigen-presenting cells that play a crucial role in the defence against pathogens. Invading pathogens are recognized by Toll-like receptors (TLRs) and receptors such as C-type lectins expressed on the surface of DCs. DCs are then able to process and present antigens, initiating primary immune response and inducing memory. They can be divided into several subsets distinguishable by specific markers and notably proteins of the C-type lectin family. Indeed, Langerhans cells (LCs) and interstitial DCs can be identified by a large panel of C-type lectins such as langerin and DC-SIGN respectively (1) (2, 3). These two lectins belong to the same family and present some strong similarity in sequence and structural organisation. Moreover, despite different specificities, they exhibit some common general recognition properties towards high mannose glycans. Indeed, they are both involved in HIV binding through specific recognition of the high mannose glycan present on the gp120 envelope protein of the virus (4). However, this shared property hides a drastically different impact in virus fate (5). While DC-SIGN can be hijacked by HIV to promote its transmission to T-Cells by

DCs (1), langerin has been reported to protect LCs from HIV-1 infection and to promote viral degradation at low viral load (6).

At high viral load or in case of a reduced expression level langerin is saturated and LCs can be infected and become efficient in HIV-1 transmission (6, 7). Indeed, co-infection with herpes simplex virus type 2 induces HIV-1 susceptibility of LCs by abrogating langerin function (8). Located in the epidermis and stratified mucosal tissue, LCs are on the frontline for interaction with pathogens, and they have been suspected to be the initial HIV targets after sexual exposure. Biological mechanisms by which HIV-1 is acquired and transmitted from mucosal site are still poorly understood. Indeed, the ability of LCs to switch from an anti-HIV-barrier status to a HIV-transmission mode, more specifically the role of langerin in this process is actively investigated.

The molecular mechanism by which langerin plays its protective role is far from being understood. Current knowledge is limited to the fact that langerin is able i) to bind HIV particles and, ii) to promote their internalization into organelles specific to LCs, the Birbeck granules (BGs). Interestingly, langerin is also the central element required for the generation of these specific organelles. Birbeck granules are pentalamellar and zippered membranes defining “tennis racket”-like structures. Langerin

accumulation is essential to generate this membrane architecture as illustrated by the induction of such cytomembrane sandwiching structures (CMS) by langerin cDNA transfection into fibroblasts or even in melanoma cell lines (3, 9). Langerin extracellular domain contains a neck region responsible for its trimeric organization and a C-terminal carbohydrate recognition domain (CRD) responsible for the  $\text{Ca}^{2+}$ -dependent lectin properties (10). In a previous study, we demonstrated the requirement of CRDs for BGs formation and suggested a macromolecular organization of langerin compatible with BG architecture (11).

To date only two point mutations leading to altered BG organization have been reported in langerin. These mutations are both located in the CRD and have been identified in human or in murine context (W264R or F244L, respectively) (12, 13). Such mutations represent an opportunity to study the relationship between BG alteration and a specific modification of langerin properties. We reconstituted these mutations in recombinant versions of langerin CRD and entire extracellular domain (ECD) to evaluate their impact on the lectin properties, notably the binding to HIV-1 envelope glycoprotein gp120.

After confirmation of the impact of mutations in BG structure, we characterized the effects of both mutations on isolated CRD and entire ECD in parallel. While the W264R mutation causes global unfolding of the CRD, the mutation of phenylalanine 241 into leucine in human langerin CRD (corresponding to F244L in the murine gene) leaves functional and structural properties of CRD globally unchanged although the whole ECD exhibits major functional and structural alterations. Indeed, the oligomerisation state of Lg ECD<sub>F241L</sub> is perturbed and consequently its global architecture as demonstrated by SAXS comparative analysis of WT and mutant forms of recombinant Lg ECD. Moreover, while the intrinsic lectin activity of the CRD is not impaired by the mutation, the avidity property of the oligomeric form is modified as shown by its weaker binding to gp120 glycan moiety. In this work we establish a direct link between a mutation in the CRD and modification of Lg ECD quaternary structure, loss of efficiency in gp120 recognition and alteration of BG architecture.

## EXPERIMENTAL PROCEDURES

*Transfection and expression of langerin constructs in COP5 cells* — Murine fibroblastic COP5 cells were cultured in RPMI 1640 medium supplemented with 10% heat-inactivated foetal bovine serum, 10 mM HEPES, 2 mM L-glutamine, and 50  $\mu\text{M}$  2-mercaptoethanol and gentamicin (80  $\mu\text{g}/\text{mL}$ ). The cells were transfected using the GeneJuice® Transfection Reagent (Merck) with cDNA encoding the various langerin expression constructs. Langerin expression was evaluated by FACS analysis 48 h after transfection and processed to electron microscopy.

*Flow cytometry analysis* — COP5 transfected cells were incubated with mAb anti-langerin 808E10 or with mAb 310F7 for intracytoplasmic staining (Dendritics). mAbs were revealed with goat anti-mouse Ig-FITC conjugated (Becton Dickinson). Intracellular staining was performed in the presence of permeabilization medium (0.3% saponin, 2% BSA). Negative controls were performed with an isotype Ig control. Fluorescence was analyzed with a FACScan® flow-cytometer (Becton Dickinson).

*Transmission electron microscopy* — After washing, isolated COP5 fibroblasts transfected with langerin cDNA were fixed with 2% glutaraldehyde in cacodylate buffer at pH 7.4 for 24 h. After rinsing in cacodylate buffer with sucrose for 24 h, cells were processed for transmission electron microscopy. They were postfixed with an aqueous solution of 1% osmium tetroxide in cacodylate buffer with sucrose and embedded in epoxy medium after dehydration through a graded series of ethanol. Ultra thin sections were stained with lead citrate and uranyl acetate and examined with a JEOL 1200EX electron microscope with acceleration voltage of 80 kV (Centre des Microstructures, Lyon University, France).

*Cloning and expression of recombinant langerin domains* — Lg CRD and Lg ECD were expressed as previously described (11, 14). Mutations in the CRD were introduced into the human langerin cDNA by PCR using synthetic oligonucleotides. The forward-primer sequence, used for F241L mutation, was 5'-GAGAGTGAGCAGGAGCTGCTGTATAAAACAGCGGG-3'. The forward-primer sequence, used for W264R human mutation, was 5'-AGGGATGGAAGGGGACCGCTCCTGGGTGGATGAC-3'. The changes were checked by DNA sequencing (Cogenics).



**Protein purification** — Lg CRD<sub>WT</sub> and Lg CRD<sub>F241L</sub> were purified using the same protocol described in (11, 14). Lg CRD<sub>W264R</sub> and all forms of Lg ECD were expressed as inclusion bodies so that a refolding step was performed before the purification procedure. The pellet obtained from a 1 L culture was resuspended into 30 mL of buffer A (25 mM Tris pH 8.0 and 150 mM NaCl) with addition of one complete EDTA-free tablet (Roche Diagnostic). Cells were lysed by a 4 min sonication with 2 s pulses separated by 10 s pauses. Inclusion bodies were isolated by centrifugation at 100 000 g for 30 min, 4°C. They were resuspended into 30 mL of buffer A with addition of 2 M urea and 1% X100 Triton and centrifuged at 100 000 g for 30 min at 4°C. Pellets were washed in 30 mL of buffer A and centrifuged again at 100 000 g for 30 min, 4°C. Inclusion bodies were then solubilised in 30 mL of buffer A with addition of 6 M Guanidine and 0.01%  $\beta$ -mercapto-ethanol. Aggregates were discarded at 100 000 g for 30 min, 4°C. The supernatant was diluted down to 2 mg/mL in the same buffer. The solution was then diluted six fold by flash-dilution into a buffer containing 25 mM Tris pH 8.0, 1.25 M NaCl and 25 mM CaCl<sub>2</sub>. A first dialysis step was performed against 25 mM Tris pH 8.0 (1 volume against 7). The resulting solution was dialyzed against buffer B (25 mM Tris pH 8.0, 150 mM NaCl and 4 mM CaCl<sub>2</sub>). For Lg ECDs, purification was performed by affinity chromatography on a mannan-agarose column (Sigma) equilibrated with buffer B and eluted in the same buffer without CaCl<sub>2</sub> but supplemented with 1 mM EDTA (buffer C). The eluted protein was dialyzed against buffer B. For refolded Lg CRD<sub>W264R</sub>, purification was performed as for Lg CRD<sub>WT</sub> using the N-terminal Strep-tag II affinity tag (14).

**Retardation assay** — Functionality Lg CRD (WT and mutants) was tested as previously described (14) on a mannan-agarose column (11 mL, Sigma).

**Crystallization, Data collection and processing** — Lg CRD<sub>F241L</sub> was crystallized by hanging drop method with initial mixing of 1  $\mu$ L of Lg CRD (5 mg/mL in buffer B) and 1  $\mu$ L of reservoir solution (100 mM MES pH 6.0; 200 mM MgCl<sub>2</sub> and 17% PEG 3350). A Lg CRD<sub>F241L</sub> crystal was cryoprotected by transfer into a solution composed of 100 mM MES pH 6.0, 200 mM MgCl<sub>2</sub> and 35%

PEG 3350 and flash-frozen in liquid nitrogen. X-ray diffraction data were collected at ID14-2 beamline (ESRF, Grenoble). 360 images were recorded with an oscillation range of 0.5° and 10 s exposure time. Data were processed using the XDS program (15). The space group is P4<sub>2</sub>, the same as that of Lg CRD. The structure was solved by molecular replacement using the MolRep program (16) and Lg CRD<sub>WT</sub> structure as model (11). The structure was refined alternating refinement cycles performed with the Refmac program (17) and manual construction steps using the Coot program (18).

Table I: Lg CRD<sub>F241L</sub> data collection and structure refinement statistics.

Data collection statistics	
Wavelength (Å)	0.933
Space group	P4 <sub>2</sub>
Unit cell parameters (Å)	a = b = 79.96 ; c = 90.42
Resolution (Å)	50- 1.4 (1.48-1.4) <sup>a</sup>
Measured reflections	834069 (130176)
Unique reflections	111136 (17644)
Completeness (%)	99.2 (98)
I/ $\sigma$ (I)	24.88 (4.7)
R <sub>merge</sub> <sup>b</sup> (%)	5.3 (32.7)
Structure refinement statistics	
Resolution (Å)	19.5-1.4
Refinement factors	
Used reflections (%)	105603/ 4.9752
R <sub>cryst</sub> <sup>c</sup>	0.1909
R <sub>free</sub> <sup>c</sup>	0.2258
Rmsd from ideality	
Bond lengths (Å)	0.009
Bond angles (de)	1.280
Average B-factor (Å <sup>2</sup> )	14.88

<sup>a</sup> Values in parentheses are for the highest resolution shell. <sup>b</sup>  $R_{\text{merge}} = \frac{\sum_h \sum_m |I_m(h) - \langle I(h) \rangle|}{\sum_h \sum_m I_m(h)}$ . <sup>c</sup>  $R_{\text{cryst}} = \frac{\sum ||F_o| - |F_c||}{\sum |F_o|}$ , and  $R_{\text{free}} = R_{\text{cryst}}$  calculated with 5% of  $F_o$  sequestered before refinement.

**Circular dichroism spectroscopy** — A Jobin Yvon CD6 circular dichroism (CD) spectropolarimeter was used. Data were recorded at room temperature in a 0.1 cm cell with a wavelength step of 1 nm and an integration time of 8 s. Each spectrum was generated by averaging three scans, after buffer correction. Protein concentration was 0.1 mg/mL for all the samples. The formula used to calculate the ellipticity is  $(33 \times \Delta A \times 100) / (l \times \text{Protein concentration} / \text{MW} \times$

(number of residues – 1)), with  $l$  = cell length (0.1 cm), (number of residues – 1) = 154 and MW = 17905.1, 17871.1 and 17875.1 Da for Lg CRD<sub>WT</sub>, Lg CRD<sub>F241L</sub> and Lg CRD<sub>W264R</sub>, respectively.

**Thermal Shift Assay** — An iQ5 real-time PCR detection system from BioRad was used. The fluorophore used for the unfolded protein detection was SYPRO orange. The temperature range was ranged from 293 to 368 K at 1 K/min. The stability of Lg CRD<sub>WT</sub>, Lg CRD<sub>F241L</sub> and Lg CRD<sub>W264R</sub> at 0.5 mg/mL was determined in 100 mM HEPES pH 7.0, 500 mM NaCl, 4 mM CaCl<sub>2</sub>, 2  $\mu$ L of 50x diluted commercial stock of SYPRO orange.

**Cross-linking coupled to MALDI-TOF mass spectrometry** — The WT and mutant Lg ECDs were cross-linked by EGS. Proteins were dialysed against 25 mM MES pH 6.0, 150 mM NaCl and 4 mM CaCl<sub>2</sub> to eliminate the Tris-HCl incompatible with the use of EGS. Proteins at 0.5 mg/mL were incubated with 5 mM EGS for 2 h at 4°C and the reaction was stopped by addition of Tris-HCl. MALDI-TOF spectra were acquired with a Bruker Daltonics Ultraflex III mass spectrometer. Native and cross-linked Lg ECD<sub>WT</sub> and Lg ECD<sub>F241L</sub> were serially diluted in 0.1% TFA and mixed on the MALDI probe with an equal volume of sinapinic acid matrix (saturated in acetonitrile/water/TFA acid 50/50/0.1, v/v/v). The TOF analyzer was set in the linear mode for protein detection. Data were analysed and smoothed using Bruker FlexAnalysis software.

**Surface plasmon resonance** — All experiments were performed on a Biacore 3000 using a functionalised CM4 chip and the corresponding reagents from Biacore. Two flow-cells were activated by 50  $\mu$ L of an EDC/NHS mixture. Flow-cell one was blocked with 50  $\mu$ L of ethanolamine and used as a control surface. Flow-cell two was functionalised at a level of 514 RU with a 10  $\mu$ g/mL solution of gp120 YU2 in sodium acetate pH 4.0 buffer. Increasing amount of Lg CRD and Lg ECD (both WT and mutant) were injected at 20  $\mu$ L/min flow rate. Lg CRDs sensorgrams were fitted with Langmuir equation using the BIAeval software. Running buffer was buffer A supplemented with 4 mM CaCl<sub>2</sub> and 0.005% P20 surfactant. Each protein was diluted into the running buffer. 20  $\mu$ L of 50 mM EDTA were injected to regenerate the chip.

**Small Angle X-ray scattering — SAXS** experiments were carried out on the SWING beamline at the SOLEIL synchrotron radiation facility (Saint-Aubin, France). The sample to detector (Aviex CCD) distance was set to 1820 mm, allowing useful data collection over the momentum transfer range  $0.015 \text{ \AA}^{-1} < q < 0.5 \text{ \AA}^{-1}$  with  $q = 4\pi \sin\theta / \lambda$ , where  $2\theta$  is the scattering angle and  $\lambda$  the wavelength of the X-rays ( $\lambda = 1.0 \text{ \AA}$ ). Preliminary experiments showed that langerin solutions always contain a large fraction of aggregates. In order to separate aggregates from isolated proteins, SAXS data were collected directly after elution through the on line size-exclusion high-performance liquid chromatography (SE\_HPLC) column available on SWING (19). Flow rate was 150  $\mu$ L/min, frame duration was 2 s and the dead time between frames 0.5 s. For each frame, the protein concentration (between 0.5 and 1 mg/mL at the top of elution peak) was estimated from UV absorption at 280 nm using a spectrometer located immediately upstream of the SAXS measuring cell. Selected identical frames corresponding to the main elution peak were averaged. A large number of frames were collected before the void volume and averaged to account for buffer scattering, which was subsequently subtracted from the averaged protein signal. The scattered intensities were put on an absolute scale using water scattering (20). Data analysis and *ab initio* shape calculations were performed using programs from the ATSAS package (21).

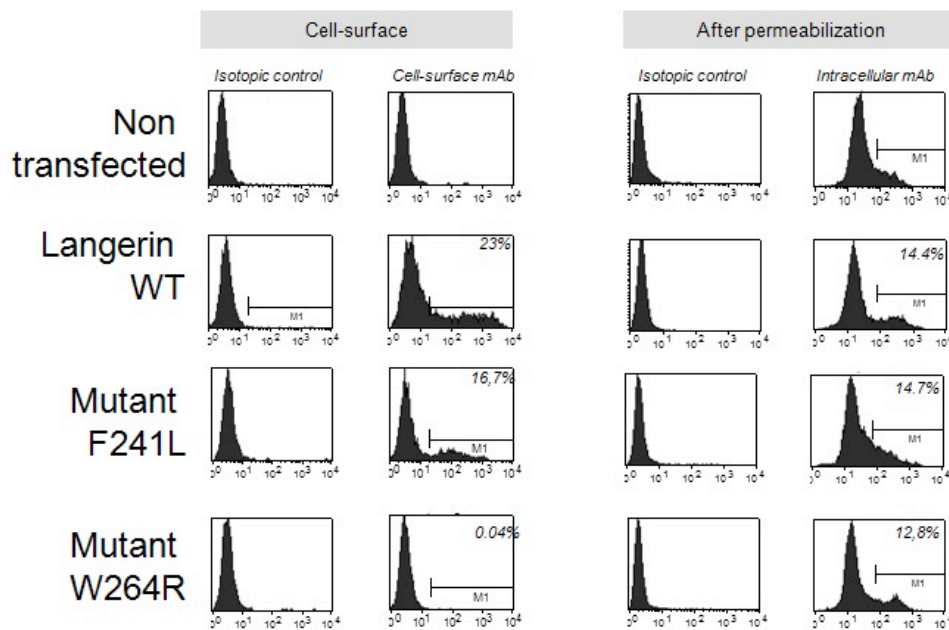


Figure 1. **FACS analysis of COP5 cells transfected or not with: Lg<sub>WT</sub>, Lg<sub>F241L</sub> and Lg<sub>W264R</sub>.** mAb anti-Langerin 808E10 was used to detect Langerin on cell surface and mAb 310F7 for intracellular staining.

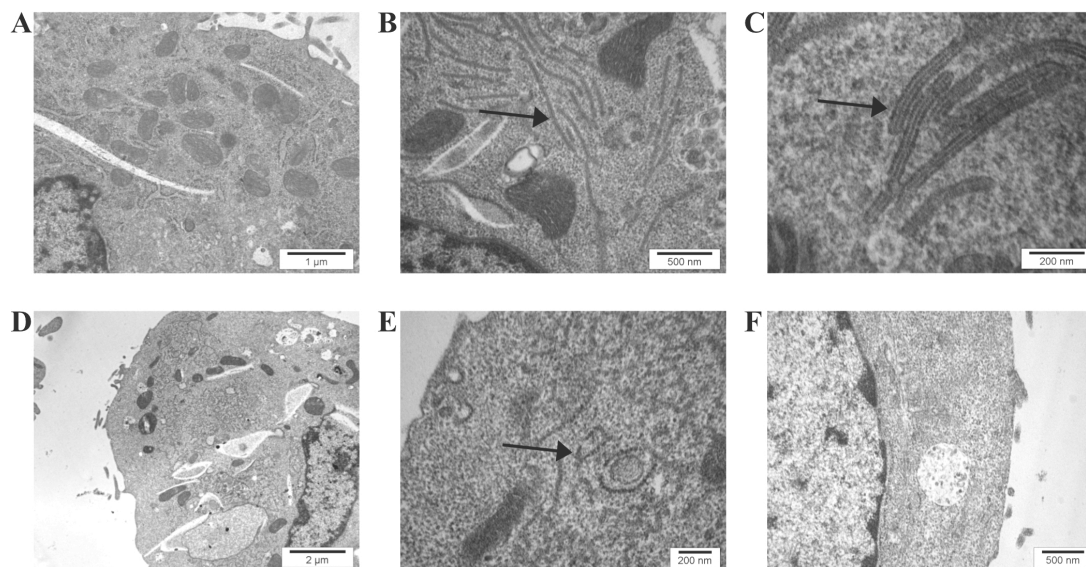


Figure 2. **Electron microscopy pictures of transfected COP5 cells.** A. Cells not transfected. B to F. Cells transfected with: Lg<sub>WT</sub> gene (B, C), Lg<sub>F241L</sub> gene (D, E) and Lg<sub>W264R</sub> gene (F). The characteristic Birbeck granule structure is shown in B and C. Arrows point to CMS (observed with Lg<sub>WT</sub>) or superimposed and thickened membranes (observed with Lg<sub>F241L</sub>). Comparison of B with E shows the structural modification of CMS caused by the F241L mutation (substitution of the double layers with an inner striation with curved and irregularly thickened double layers without inner electron-dense bodies) while panel F shows the loss of CMS with the W264R mutation.

## RESULTS AND DISCUSSION

### Impact of langerin mutations on Birbeck granule structure

Langerin gene transfection in COP5 fibroblasts leads to the formation of cytomembrane sandwiching structures (CMS) typical of Birbeck granules (BGs) (3).

Mutation-induced alteration of Birbeck granule structure was reported for mutation F244L in the murine langerin gene transfected-COP5 (13) and for mutation W264R isolated from a healthy person whose LCs lack BG altogether (12). Considering the difference in gene origin and cellular environment of these two reported mutations, we decided to analyse those two mutations within the same genetic and cellular

context, *i.e.* the human gene transfected in the previously described COP5 fibroblast model (Fig. 1). First, using flow cytometry, we controlled the cell surface expression of langerin on COP5 transfected cells using mAb anti-Lg clone 808E10 directed against a CRD epitope, and intracytoplasmic expression using mAb anti-Lg clone 310F7 which recognizes an intracytoplasmic epitope. As shown in Fig. 1, although transfection level reaches only 12 to 20%, expression of langerin is clearly observed using the two mAbs except for the mutation W264R form where DCGM4 staining on cell surface is not observed probably due to an epitope disappearance as following characterization suggests.

Typical CMS were observed following transfection of the WT human langerin gene (Fig. 2B-C). Phenylalanine 244 in murine langerin is conserved in human as phenylalanine 241 as shown by sequence alignment (13). The transfection of Lg ECD<sub>F241L</sub> gene leads to the formation of superimposed and thickened membranes structures different from typical CMS of BG, as previously described with the mouse gene (Fig. 2D-E). We also confirm the total loss of BG associated with W264R mutation already reported (Fig. 2F).

These two mutations are located within the carbohydrate recognition domain reinforcing the essential role of this domain in BG formation as already suggested (11). To understand from a molecular point of view the effect of these mutations, different recombinant carbohydrate recognition domains containing those mutations as well as full-length extracellular domains (ECD), have been produced and characterised.

### **Characterisation of mutation effects on langerin CRD**

*Langerin wild type and mutants CRD production and purification* — For the Lg CRDs purification step, we used the Strep-tag II as described in (14). The soluble expression of the Lg CRD, in the *E. coli* periplasm, is a first indication of a proper folding of Lg CRDs. This is confirmed by the similar elution profiles of Lg CRD<sub>wt</sub> and Lg CRD<sub>F241L</sub> of the soluble fraction (data not shown). In contrast, Lg CRD<sub>W264R</sub> was not present in the soluble fraction and formed

inclusion bodies (data not shown). As reported for CRDs of other C-type lectin such as DC-SIGN (22, 23), a refolding step yielded a soluble form of Lg CRD<sub>W264R</sub>.

#### *Mannan binding properties of mutants* —

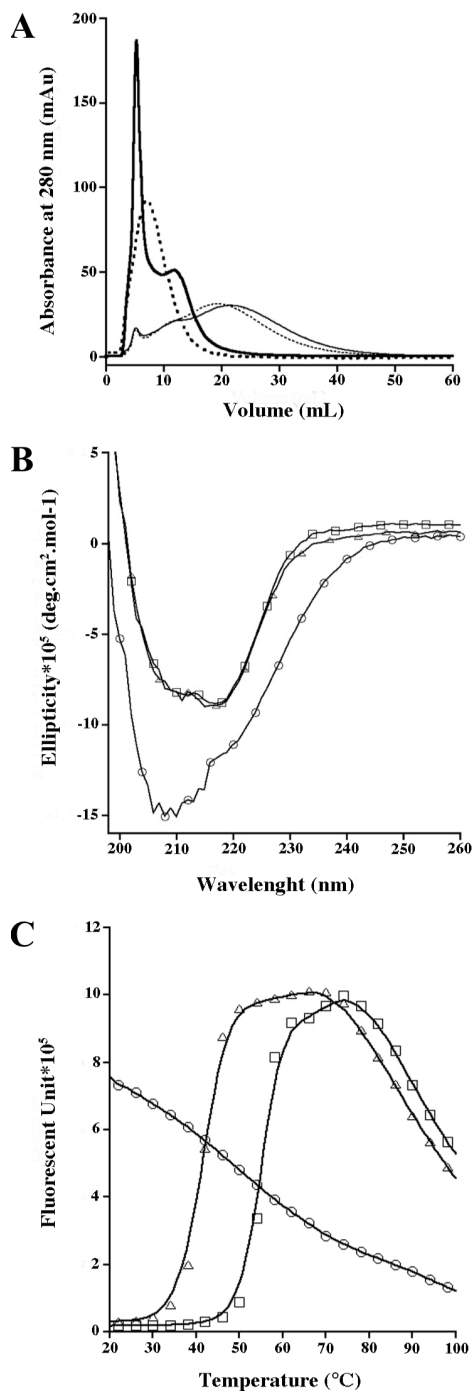
Functionality of the various Lg CRD mutants was evaluated with the wild type as positive control, using the delayed elution of proteins on a mannan-agarose column as a marker of their affinity for the resin (11). Bovine serum albumin, used as negative control, eluted in the void volume of the column while Lg CRD<sub>F241L</sub> eluted with a delay comparable to that observed with Lg CRD<sub>wt</sub> (Fig. 3A). The F241L mutation does not significantly affect the capacity of the CRD to interact with the mannan. In contrast, Lg CRD<sub>W264R</sub> eluted in the void volume as the negative control, suggesting that despite the isolation of a soluble form after refolding, this one is not functional.

#### *Analysis of gp120 binding properties* —

Binding of Lg CRDs to the envelope glycoprotein gp120 of HIV-1 has been studied using surface plasmon resonance (Fig. 4). We observed very similar interaction signals for both Lg CRD<sub>WT</sub> and Lg CRD<sub>F241L</sub> with  $K_d$  values of 24  $\mu$ M and 5  $\mu$ M respectively derived from a Langmuir analysis ( $\chi^2$  values around 0.5). The CRD harbouring the F241L mutation appears to be fully functional as regards gp120 recognition.

#### *Structural integrity of WT and mutants Lg-CRD* —

The secondary structure composition of the three CRDs has been evaluated by CD (Fig. 3B). We observed a perfect superposition of the Lg CRD<sub>F241L</sub> and Lg CRD<sub>wt</sub> spectra indicative of identical secondary structures. In contrast, the CD pattern of Lg-CRD<sub>W264R</sub> was markedly different from that of Lg CRD<sub>wt</sub>. Structural stability has been investigated using thermal shift assay. Misfolding of Lg CRD<sub>W264R</sub> is confirmed by the strong fluorescence signal observed at 20°C indicating that hydrophobic regions are already exposed to the solvent (Fig. 3C). The fluorescence decrease at higher temperature is due to protein aggregation. In view of the location of W264 within a central hydrophobic core of the protein, it could be anticipated that its replacement by an arginine would cause the observed CRD global unfolding (Fig. 5B). While Lg CRD<sub>F241L</sub> behaves as a well-folded protein, with a clear melting transition, it exhibits a  $T_m$  10°C lower than that of the wild-type protein (Fig. 3C). This reduced stability is the first noticeable effect of the mutation.



**Figure 3. Functional and structural characterization of Lg CRDs.** A. Delay experiment of Lg CRDs and BSA in Mannan-Agarose affinity column. BSA, used as a negative control (dotted bold line), Lg CRD<sub>WT</sub> (thin line), Lg CRD<sub>F241L</sub> (dotted thin line) and Lg CRD<sub>W264R</sub> (thick line). B. CD spectra of Lg CRD<sub>WT</sub> (□), Lg CRD<sub>F241L</sub> (△) and Lg CRD<sub>W264R</sub> (○). C. Thermal shift assay of Lg CRD<sub>WT</sub> (□), Lg CRD<sub>F241L</sub> (△) and Lg CRD<sub>W264R</sub> (○) as a function of the temperature.

*X-ray structure of Lg CRD<sub>F241L</sub>* —To explain the stability decrease in F241L mutant, we crystallised and solved the structure of Lg CRD<sub>F241L</sub> (Table I, PDB code: 4ak8). Superimposition of Lg CRD<sub>F241L</sub> structure over that of Lg CRD<sub>WT</sub> structure, previously solved by us and others (11, 24), shows that the replacement of the phenylalanine by a leucine causes the loss of a  $\pi$  orbital stacking between the F241 and the F202 from  $\alpha$ 2 helix and  $\beta$ 1 strand (Fig. 5A-C-D). This interaction may play a role in anchoring the first N-terminal  $\beta$ -sheet onto the rest of the CRD and its loss may explain the reduced thermal stability of the F241L mutant reported above. Moreover, the mutation induces a large movement of K200 that fills the space left empty by the loss of the phenylalanine ring (Fig. 5D). As a result, K200 establishes a new hydrogen bond with T245 in Lg CRD<sub>F241L</sub>.

The net contribution of this additional H-bond to the stability is lower than the initial hydrophobic cluster centred on F241. H-bond being more sensitive to temperature increase, this explains the lower stability achieved. However, these modifications are essentially local, do not significantly impair the recognition properties of gp120, and do not provide an easy explanation for the deep disorganisation of the Birbeck granule super-structure (Fig. 2).

Feinberg *et al* recently published a structure of a trimeric truncated form of the langerin extracellular domain (25). In this trimer, F241 is found at the interface between neighbouring monomers so that its mutation is likely to destabilize interactions within the trimer. To clarify this issue, we decided to study this mutation in the context of the whole extracellular domain.



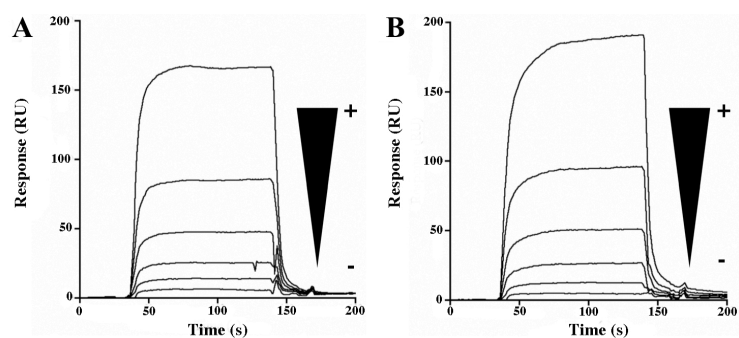


Figure 4. **Sensorgrams of the interaction of Lg CRDs with gp120.** A. Lg CRD<sub>WT</sub>. B. Lg CRD<sub>F241L</sub>. Both proteins were injected in series of twofold dilutions over the concentration range [12.5  $\mu$ M, 400  $\mu$ M].

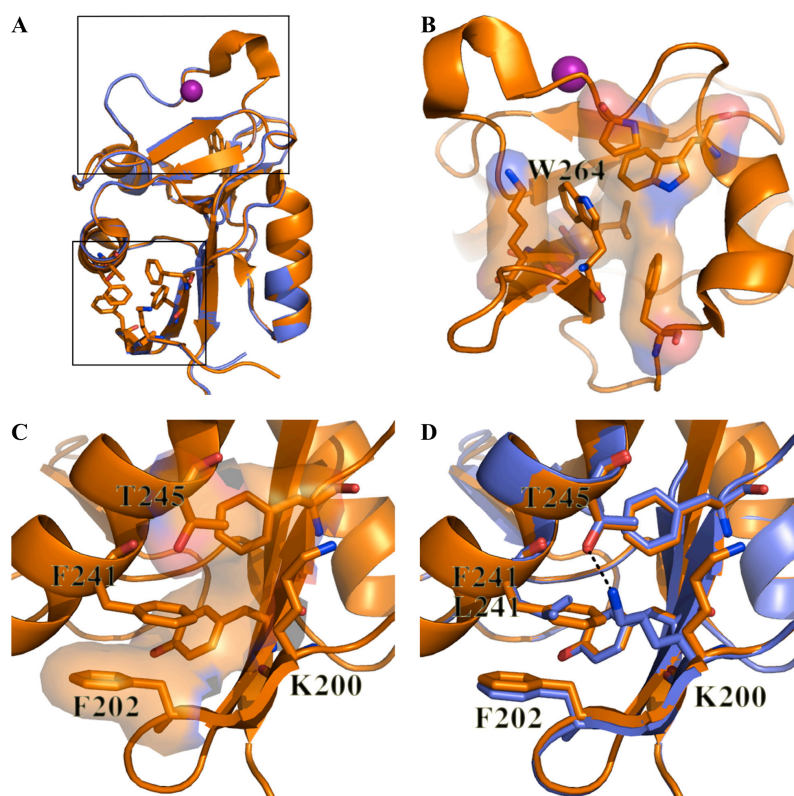


Figure 5. **Crystallographic structures of Lg CRD<sub>WT</sub> (orange) and Lg CRD<sub>F241L</sub> (blue).** A. Superimposition of both structures. B. Focus on W264 hydrophobic cluster in Lg CRD<sub>WT</sub> structure. C. Focus on F241 and F202  $\pi$  orbital stacking in Lg CRD<sub>WT</sub> structure. D. Same focus as C with both superimposed structures.

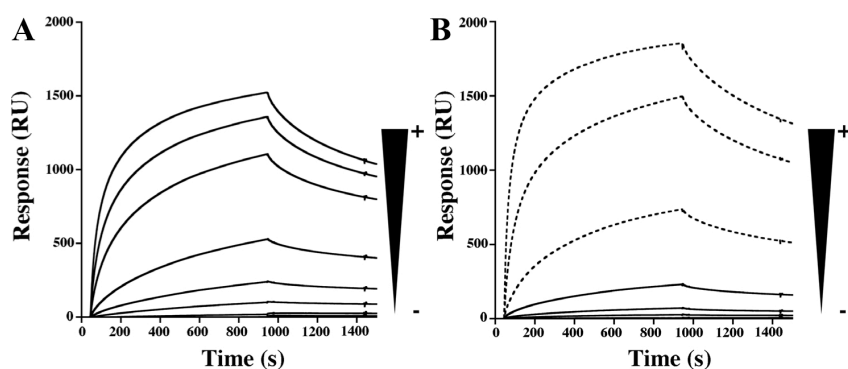


Figure 6. **Sensorgrams of the interaction of Lg ECDs with gp120.** A. Lg ECD<sub>WT</sub>; concentrations used are 5, 10, 50, 100, 200, 500, 750 and 1000 nM. B. Lg ECD<sub>F241L</sub>; black lines correspond to concentrations of 62.5, 250, 500 and 1000 nM and dotted lines correspond to additional concentrations used, 2, 4 and 6.6  $\mu$ M.

### Characterisation of F241L mutation effects on langerin ECD

*Lg ECD<sub>WT</sub> and Lg ECD<sub>F241L</sub> production and purification* — Recombinant langerin ECD is produced in inclusion bodies in bacteria and must therefore be refolded before protein purification, whereupon functional oligomers are sorted by affinity purification onto mannan agarose column. Both Lg ECD<sub>WT</sub> and Lg ECD<sub>F241L</sub> bound to the column and required EDTA for elution in contrast to the CRDs that were just delayed. This is due to an avidity effect, thereby revealing the existence of oligomeric forms for each protein (data not shown).

*Alteration of gp120 binding properties of Lg ECD<sub>F241L</sub>* — Titration experiments of both Lg ECD<sub>WT</sub> and Lg ECD<sub>F241L</sub> were performed on an immobilised gp120 surface as described above for CRDs. Both ECDs are able to recognise high mannose glycans from gp120 but Lg ECD<sub>F241L</sub> affinity is 7 times lower than that of the WT (Fig. 6). Since, Lg CRD<sub>F241L</sub> affinity was slightly better, if at all different, from that of Lg CRD<sub>WT</sub>, the weaker affinity of Lg ECD<sub>F241L</sub> implies a reduction in the avidity properties of the mutant, due to the modification of the CRDs arrangement or even of the oligomeric state of the protein.

*Cross-linking coupled to MALDI-MS of Lg ECD<sub>WT</sub> and Lg ECD<sub>F241L</sub>* — To address this question; we first performed cross-linking experiments coupled with mass spectrometry. While results show the expected trimer of WT protein, a dimer is clearly observed for Lg ECD<sub>F241L</sub> (Fig. 7). This prompted us to undertake a SAXS study in order to investigate into more detail the oligomeric state and conformation of both ECDs.

### SAXS analysis of Lg ECD<sub>WT</sub> and Lg ECD<sub>F241L</sub>

Preliminary experiments showed that concentrated langerin solutions always contain a large fraction of aggregates. In order to separate aggregates from isolated proteins, SAXS data were collected after elution through the on line SE-HPLC column available on the SWING beamline at the Synchrotron SOLEIL (19). Elution profiles measured with the UV spectrometer ( $\lambda=280$  nm) are displayed in Fig. 8 for both

WT and mutant proteins. In the case of WT protein the profile exhibits a single narrow peak beyond that corresponding to aggregates. In the case of the mutant protein, a clear shoulder is seen on the rising edge of the main elution peak that is markedly shifted towards longer elution times.

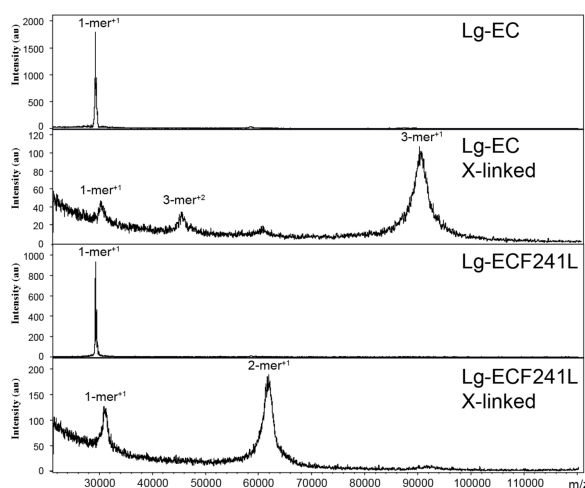


Figure 7. **Oligomeric state of Lg ECD assessed by cross-linking and MALDI-MS analysis.** Top to bottom panel correspond to MALDI-MS spectra of Lg ECD<sub>WT</sub>, Lg ECD<sub>WT</sub> cross-linked with EGS, Lg ECD<sub>F241L</sub>, Lg ECD<sub>F241L</sub> cross-linked with EGS, respectively.

*Lg ECD<sub>WT</sub>* — Scattering patterns recorded around the maximum of the elution peak are identical, so that 20 frames could be averaged to yield a scattering curve amenable to further analysis. The estimate of the molecular mass derived from  $I(0)/c$  value is of the order of 90 kDa, in good agreement with the molecular mass of a trimer (MW = 88 219 Da). The values of the radius of gyration ( $R_g$ ) and of the maximal diameter ( $D_{max}$ ) derived from the data are given in Table II. Previously, we proposed a structural model of the ECD based on the results of hydrodynamic studies (11). We modified this model by incorporating the recent crystal structure of the trimeric head of langerin containing the CRD and one coiled-coil repeat (pdb 3KQG) (25) (Fig. 9A). We calculated the scattering pattern  $I(q)$  of the resulting model together with the corresponding distance distribution function  $p(r)$ . Comparison of the  $R_g$  and  $D_{max}$  values of the model with experimental values (Table 2) and of its  $p(r)$  with the experimental one (Fig. 9B) clearly indicate that the protein neck in solution does not adopt the straight conformation proposed in our initial model (11).

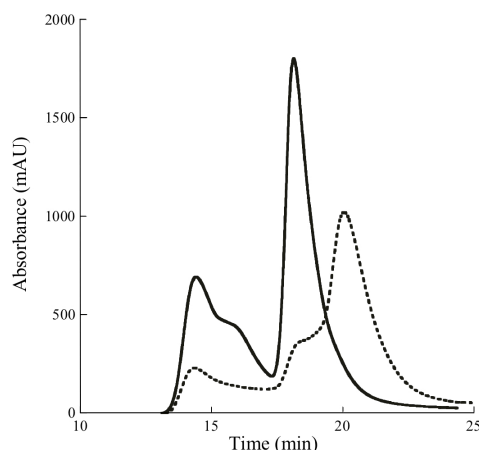


Figure 8. SAXS-coupled SE-HPLC elution profiles of Lg ECDs. Lg ECD<sub>WT</sub> (continuous line) and Lg ECD<sub>F241L</sub> (dashed line).

These conclusions are confirmed by *ab initio* modelling of the protein from the experimental scattering pattern using the program GASBOR that describes the protein as a chain of dummy residues (26). The resulting conformations all exhibit a broader extremity (the “head”) followed by a curved tail. Two orthogonal views of a typical model are shown in Fig. 9C. The structure of the trimeric head (pdb 3KQG) yields a reasonable fit when superimposed over the head part of our model. Furthermore, the similarity between calculated and experimental  $p(r)$  profiles in the range  $30 \text{ \AA} < r < 50 \text{ \AA}$  suggests that the arrangement of the three CRDs in solution is close to that observed in the crystal, although a slightly more compact arrangement of the head cannot be ruled out. This agreement points towards a rather rigid organisation of the polypeptide chains in the intermediate region between neck and CRDs.

In the absence of complementary information, our observations regarding the curved neck can be interpreted in several ways: i) the coiled-coil presents one or several kinks; ii) the coiled-coil exhibits some flexibility around the kinks; iii) the neck may not adopt a coiled-coil conformation along its entire length, in particular in the N-terminal region.

We have tried to fit the experimental scattering pattern by introducing two kinks in the regions where coiled-coil prediction programs suggest an interruption of the coiled-coil (arrows in Fig. 9D), while leaving the three resulting domains free to move under

distance restraints using the program SASREF (27).

An example of resulting conformation is shown in Fig. 9E. No perfect fit is ever obtained suggesting that the protein does not adopt a unique conformation and/or that the coiled-coil is not as extended as proposed in the model, with individual chains fraying out at the N-terminal end. Here again, langerin appears to be at variance with other C-type lectins. In DC-SIGN and DC-SIGNR, necks are essentially straight and rather rigid but comprise almost strictly conserved repeated units leading to a very homogeneous neck structure (23, 28) while in langerin, although coiled-coil motifs can be recognized along the neck, their sequence variability could account for the bent SAXS derived envelope suggesting an increased neck flexibility, the existence of gaps in the coiled-coil around kinks and finally a less extended Lg-ECD than proposed initially (11).

**Lg ECD<sub>F241L</sub>** — The elution profile shown in Fig. 8 exhibits a pronounced shoulder at an elution time similar to that of the WT trimer followed by the main peak. Each scattering pattern along the elution profile was individually analyzed. Fig. 9F presents the variation along the elution profile of the protein concentration  $c$  (measured using OD at 280 nm), the radius of gyration  $R_g$ , and  $I(0)/c$  that is proportional to the mass of the scattering object.  $I(0)/c$  values show that the scattering species corresponding to the shoulder and the top of the elution peak are compatible with trimers and dimers respectively.

Identical frames within the shoulder and around the top of the peak (5 and 18 frames respectively) were averaged so as to improve the statistical quality for further analysis and modelling of both trimer and dimer.

The trimer of the mutant protein is different from that of wild-type langerin as indicated by the smaller  $R_g$  value and by the profile of the distance distribution function that is not compatible with the semi-rigid neck following a head domain observed for the wild-type trimer (Fig. 9G). *Ab initio* analysis produces a great variety of elongated shapes with no typical shape emerging from the results. This may be due to the fact that the trimer solution is not perfectly monodispersed but may contain a small proportion of other objects, as suggested by the short duration of the plateau in the profile of  $R_g$  value across the peak of trimer shown in Fig. 9F.

The distance distribution function of the dimer is very different from that of WT langerin (Fig. 9I). The maximal extension is much shorter,



indicating that the dimer cannot adopt a conformation analogue to that of the WT trimer with a head part followed by a semi-rigid dimeric coiled-coil. However, the asymmetric  $p(r)$  is characteristic of an elongated object, which is confirmed by *ab initio* analysis using GASBOR (26) (Fig. 9H). The value of the dimer maximal extension is 140 Å, much shorter than the length of one helix from the neck in straight conformation (~200 Å). Therefore helices have to fold over, possibly interacting with one another or with the CRDs. At that point, more information is required to determine the mutual arrangement of the two protomers.

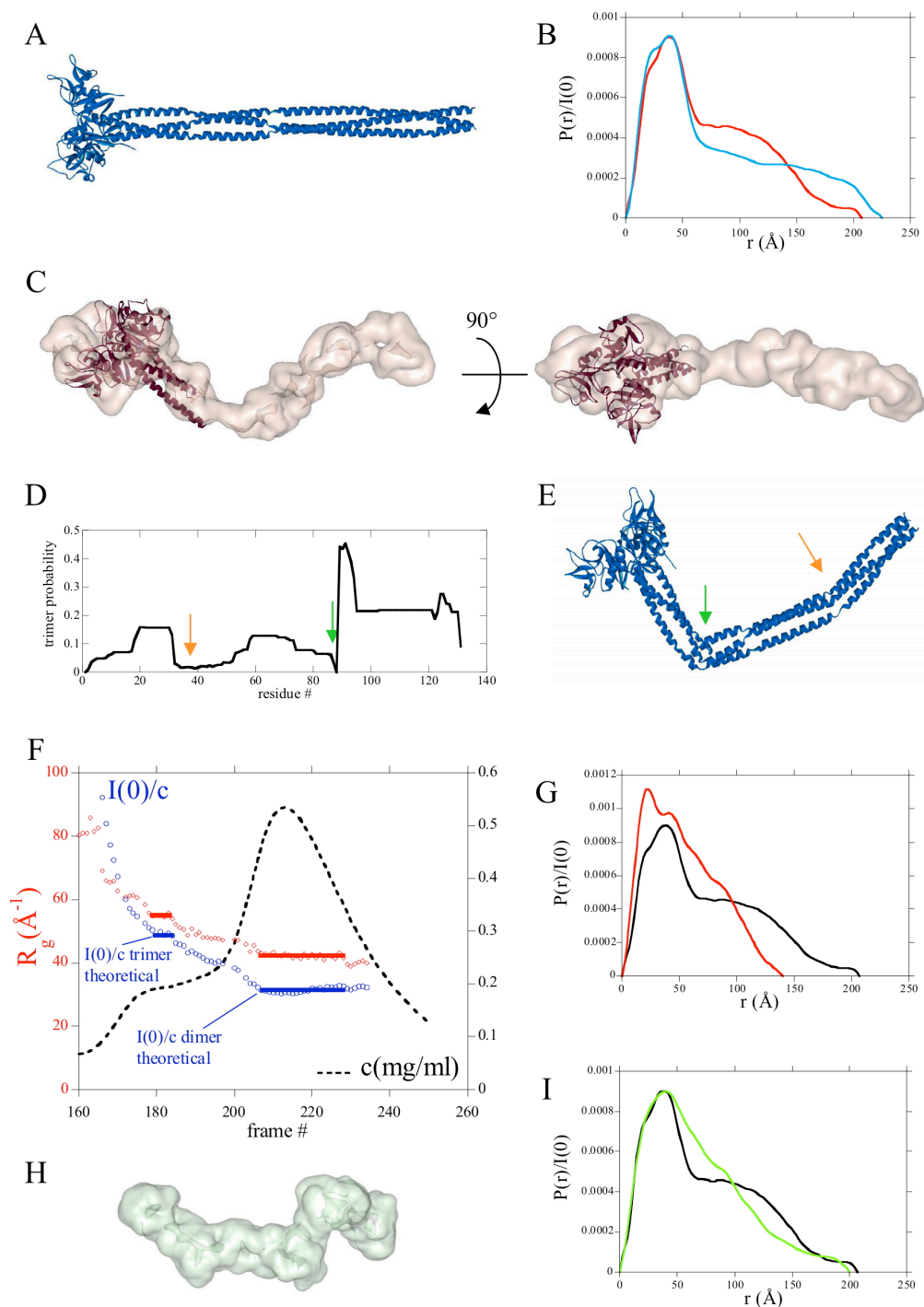
In the wild type trimer, the CRD surface containing F241 residue lies within the interface between two protomers (Fig. 10C-D) where the bottom of one CRD (containing F241) interacts with the C-terminal end of the coiled-coil neck of the neighbouring monomer (Fig. 10D). Fig. 10B, where the major part of CRDs has been omitted for the sake of clarity, shows that F241 residue from each protomer interacts directly with Q197 of the neighbouring protomer. Beyond this direct interaction between protomers, the bulky side-chain of F241 is located in the centre of the CRD/CRD interface (Fig. 10D) so that the F241L mutation, together with the already mentioned relocation of K200 (Fig. 5D) modifies the CRD surface involved in this contact (Fig. 10D). Our data globally suggest that phenylalanine in position 241 is crucial for the interface stability required for trimer formation. Indeed, the neck domain by itself cannot ensure trimer formation. Actually, analysis of coiled-coil propensity using the program Multicoil (29) that distinguishes between dimeric and trimeric coiled-coils unambiguously favours dimer formation. In langerin, trimer stabilisation appears to result from synergistic contributions from both CRD

and neck domains. The apparently minor perturbation associated with the substitution of the phenyl ring by the aliphatic side chain of leucine and lysine 200 reorientation is enough to tip the balance towards the dimeric coiled coil as observed by MALDI-MS study of cross-linked ECDs and by SAXS.

In a previous study, we proposed a molecular mechanism for the membrane zippering leading to Birbeck granule formation, in which CRDs within facing membranes could assemble and form a higher order structure at the centre of the CMS (11). This proposal was mainly supported by electron microscopy observations, notably the disappearance of CMS in COP5 transfected with langerin lacking the entire carbohydrate domain. In this study, we analysed two single mutations within the CRD. W264R mutation, similarly to CRD deletion, causes the disappearance of BG by preventing proper CRD folding. More subtle but far-reaching are the consequences of the rather conservative substitution of hydrophobic leucine for hydrophobic phenylalanine. Lg CRD<sub>F241L</sub> is still properly folded and fully functional, structural changes are few and local but located at an interface between neighbouring CRDs. This interaction is perturbed enough to switch the global balance between trimer and dimer formation finally leading to the deep alteration of CMS organization and to the assembly of irregularly thickened membrane structures observed in COP5 transfected cells (Fig. 2E). The spatial arrangement of CRDs within the trimeric head of langerin is thus an absolute requirement for a proper BG architecture. Finally, in addition to BG formation, the affinity of Lg ECD<sub>F241L</sub> for high mannose glycoconjugates in gp120 is severely reduced, showing that the specific spatial arrangement of CRDs is also key to optimal recognition of the HIV envelope glycoprotein.

Table II: structural parameters derived from SAXS data.

Molecule	$R_g$	$D_{max}$
Lg ECD <sub>WT</sub>	$60.5 \pm 1 \text{ Å}$	$207 \pm 10 \text{ Å}$
Lg ECD <sub>F241L</sub> , trimer	$55.9 \pm 1 \text{ Å}$	$200 \pm 10 \text{ Å}$
Lg ECD <sub>F241L</sub> , dimer	$42.9 \pm 1 \text{ Å}$	$140 \pm 10 \text{ Å}$
Lg ECD model	$69 \text{ Å}$	$225 \text{ Å}$



**Figure 9. SAXS analysis.** A. Model of Lg ECD<sub>WT</sub> trimer from (*11*) after incorporation of the trimeric head and proximal coiled-coil region from 3KQG. B. Comparison of the distance distribution function of Lg ECD<sub>WT</sub> model shown in A (blue line), with that derived from the experimental curve for Lg ECD<sub>WT</sub> (red curve). C. Typical envelope for Lg ECD<sub>WT</sub> shown in two orthogonal views with the crystal structure of Feinberg et al. (pdb 3KQG) manually superimposed. D. Probability of a trimeric coiled-coil along the neck sequence predicted by Multicoil (29). Two colour arrows point to the two regions of lowest probability. E. Model obtained by rigid body refinement allowing flexibility at the two points along the neck identified in D. F. Variation along the SEC elution profile of the protein concentration  $c$  (black dashed line), radius of gyration  $R_g$  (red diamonds) and  $I(0)/c$  (blue squares). Each frame corresponds to a 2s SAXS pattern. G. Comparison of the distance distribution functions of the WT trimer (black line) and of the trimer of F241L mutant (red line). H. Typical envelope of the dimer obtained using the program GASBOR. I. Comparison of the distance distribution functions of the WT trimer (black line) and of the dimer of F241L mutant (green line).

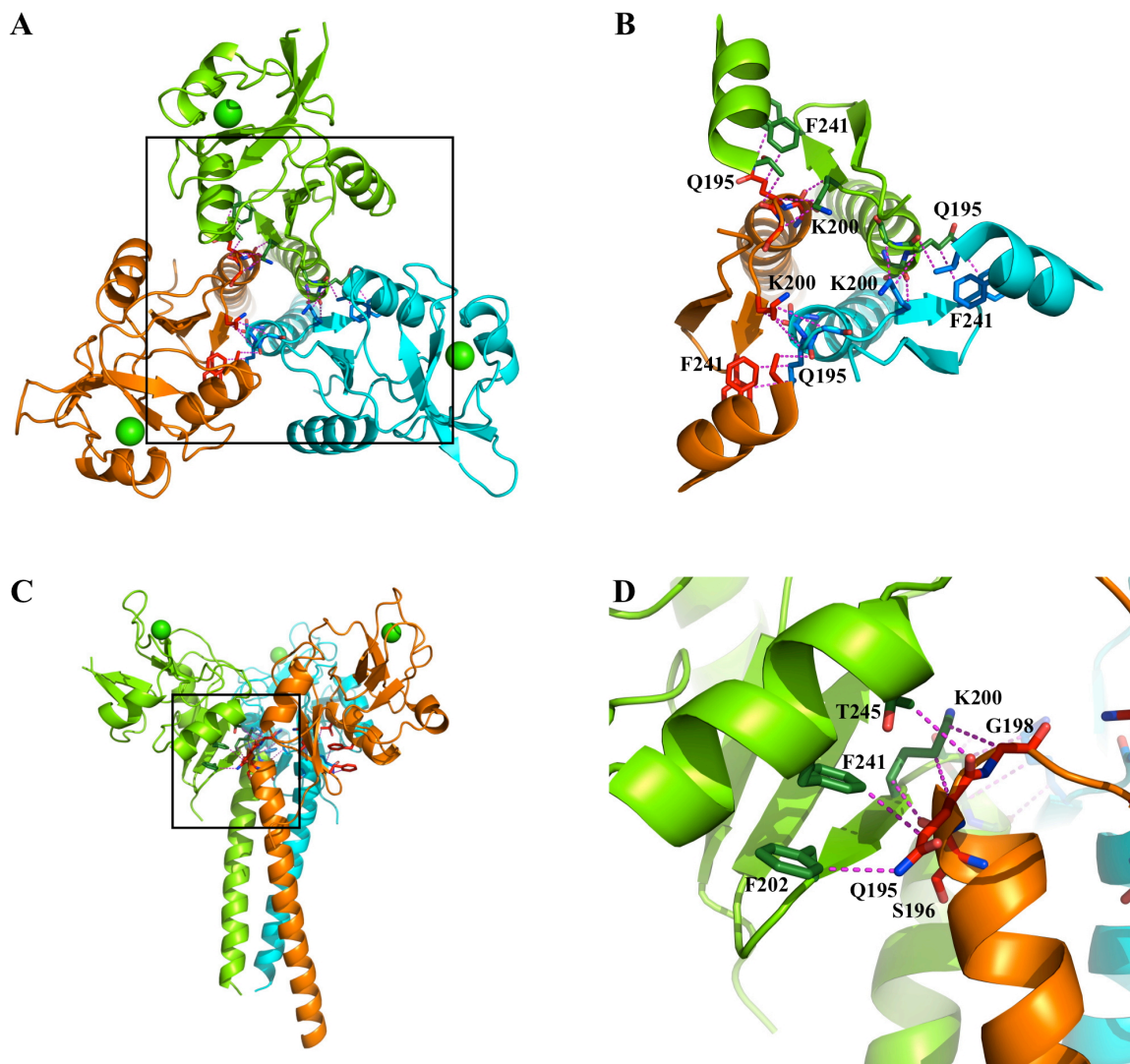


Figure 10. **F241 and K200 are central residues of the interface between CRD of one protomer and C-terminal neck end of the neighbouring one.** A. Top view. B. Top view with a zoom over the square in panel B. C. Lateral view. D. Lateral view with a zoom over the square in panel D.

## REFERENCES

1. Geijtenbeek, T. B., Kwon, D. S., Torensma, R., van Vliet, S. J., van Duijnhoven, G. C., Middel, J., Cornelissen, I. L., Nottet, H. S., KewalRamani, V. N., Littman, D. R., Figdor, C. G., and van Kooyk, Y. (2000) DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells, *Cell* 100, 587-597.
2. Valladeau, J., Duvert-Frances, V., Pin, J. J., Dezutter-Dambuyant, C., Vincent, C., Massacrier, C., Vincent, J., Yoneda, K., Banchereau, J., Caux, C., Davoust, J., and Saeland, S. (1999) The monoclonal antibody DCGM4 recognizes Langerin, a protein specific of Langerhans cells, and is rapidly internalized from the cell surface, *Eur J Immunol* 29, 2695-2704.
3. Valladeau, J., Ravel, O., Dezutter-Dambuyant, C., Moore, K., Kleijmeer, M., Liu, Y., Duvert-Frances, V., Vincent, C., Schmitt, D., Davoust, J., Caux, C., Lebecque, S., and Saeland, S. (2000) Langerin, a novel C-type lectin specific to Langerhans cells, is an endocytic receptor that induces the formation of Birbeck granules, *Immunity* 12, 71-81.

4. Turville, S. G., Cameron, P. U., Handley, A., Lin, G., Pohlmann, S., Doms, R. W., and Cunningham, A. L. (2002) Diversity of receptors binding HIV on dendritic cell subsets, *Nat Immunol* 3, 975-983.
5. de Witte, L., Nabatov, A., and Geijtenbeek, T. B. (2008) Distinct roles for DC-SIGN+-dendritic cells and Langerhans cells in HIV-1 transmission, *Trends Mol Med* 14, 12-19.
6. de Witte, L., Nabatov, A., Pion, M., Fluitsma, D., de Jong, M. A., de Gruijl, T., Piguet, V., van Kooyk, Y., and Geijtenbeek, T. B. (2007) Langerin is a natural barrier to HIV-1 transmission by Langerhans cells, *Nat Med* 13, 367-371.
7. de Jong, M. A., and Geijtenbeek, T. B. (2010) Langerhans cells in innate defense against pathogens, *Trends Immunol* 31, 452-459.
8. de Jong, M. A. W. P., de Witte, L., Taylor, M. E., and Geijtenbeek, T. B. H. (2010) Herpes Simplex Virus Type 2 Enhances HIV-1 Susceptibility by Affecting Langerhans Cell Function, *The Journal of Immunology*.
9. McDermott, R., Bausinger, H., Fricker, D., Spehner, D., Proamer, F., Lipsker, D., Cazenave, J. P., Goud, B., De La Salle, H., Salamero, J., and Hanau, D. (2004) Reproduction of Langerin/CD207 traffic and Birbeck granule formation in a human cell line model, *J Invest Dermatol* 123, 72-77.
10. Stambach, N. S., and Taylor, M. E. (2003) Characterization of carbohydrate recognition by langerin, a C-type lectin of Langerhans cells, *Glycobiology* 13, 401-410.
11. Thépaut, M., Valladeau, J., Nurisso, A., Kahn, R., Arnou, B., Vivès, C., Sealand, S., Ebel, C., Monnier, C., Dezutter-Dambuyant, C., Imberty, A., and Fieschi, F. (2009) Structural Studies of Langerin and Birbeck Granules: a Macromolecular Organization Model, *Biochemistry* 48, 2684-2698.
12. Verdijk, P., Dijkman, R., Plasmeijer, E. I., Mulder, A. A., Zoutman, W. H., Mieke Mommaas, A., and Tensen, C. P. (2005) A lack of Birbeck granules in Langerhans cells is associated with a naturally occurring point mutation in the human Langerin gene, *J Invest Dermatol* 124, 714-717.
13. Valladeau, J., Clair-Moninot, V., Dezutter-Dambuyant, C., Pin, J. J., Kissenpfennig, A., Mattei, M. G., Ait-Yahia, S., Bates, E. E., Malissen, B., Koch, F., Fossiez, F., Romani, N., Lebecque, S., and Saeland, S. (2002) Identification of mouse langerin/CD207 in Langerhans cells and some dendritic cells of lymphoid tissues, *J Immunol* 168, 782-792.
14. Thépaut, M., Vivès, C., Pompidor, G., Kahn, R., and Fieschi, F. (2008) Overproduction, purification and preliminary crystallographic analysis of the Carbohydrate Recognition Domain of human Langerin, *Acta Crystallogr F* 64, 115-118.
15. Kabsch, W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants, *J. Appl. Cryst.* 26, 795-800.
16. Vagin, A., and Teplyakov, A. Molecular replacement with MOLREP, *Acta Crystallogr D Biol Crystallogr* 66, 22-25.
17. Collaborative Computational Project, N. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* 50, 760-763.
18. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr D Biol Crystallogr* 60, 2126-2132.
19. David, G., and Pérez, J. (2009) Combined sampler robot and high-performance liquid chromatography: a fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline, *J. Appl. Cryst.* 42, 892-900.
20. Orthaber, D., Bergmann, A., and Glatter, O. (2000) SAXS experiments on absolute scale with Kratky systems using water as a secondary standard, *J. Appl. Cryst.* 33, 218-225.
21. Konarev, P. V., Petoukhov, M. V., Volkov, V. V., and Svergun, D. I. (2006) ATSAS 2.1, a program package for small-angle scattering data analysis, *J. Appl. Cryst.* 39, 277-286.
22. Mitchell, D. A., Fadden, A. J., and Drickamer, K. (2001) A novel mechanism of carbohydrate recognition by the C-type lectins DC-SIGN and DC-SIGNR. Subunit organization and binding to multivalent ligands, *J Biol Chem* 276, 28939-28945.
23. Tabarani, G., Thepaut, M., Stroebel, D., Ebel, C., Vives, C., Vachette, P., Durand, D., and Fieschi, F. (2009) DC-SIGN Neck Domain Is a pH-sensor Controlling Oligomerization:

- SAXS AND HYDRODYNAMIC STUDIES OF EXTRACELLULAR DOMAIN, *J Biol Chem* 284, 21229-21240.
24. Chatwell, L., Holla, A., Kaufer, B. B., and Skerra, A. (2008) The carbohydrate recognition domain of Langerin reveals high structural similarity with the one of DC-SIGN but an additional, calcium-independent sugar-binding site, *Mol Immunol* 45, 1981-1994.
  25. Feinberg, H., Powlesland, A. S., Taylor, M. E., and Weis, W. I. (2010) Trimeric structure of langerin, *J Biol Chem* 285, 13285-13293.
  26. Svergun, D. I., Petoukhov, M. V., and Koch, M. H. (2001) Determination of domain structure of proteins from X-ray solution scattering, *Biophys J* 80, 2946-2953.
  27. Petoukhov, M. V., and Svergun, D. I. (2005) Global rigid body modelling of macromolecular complexes against small-angle scattering data, *Biophys J*.
  28. Feinberg, H., Tso, C. K., Taylor, M. E., Drickamer, K., and Weis, W. I. (2009) Segmented helical structure of the neck region of the glycan-binding receptor DC-SIGNR, *J Mol Biol* 394, 613-620.
  29. Wolf, E., Kim, P. S., and Berger, B. (1997) MultiCoil: a program for predicting two- and three-stranded coiled coils, *Protein Sci* 6, 1179-1189.

*Acknowledgments* — We are very grateful to P. Roblin for expert assistance during measurements at the SWING beamline (Synchrotron Soleil) and to J. Pérez for insightful discussions. We also acknowledge the Centre des Microstructures from Lyon University for access to electron microscopes.

This article is dedicated to our beloved colleague, Richard Kahn who passed away in October 2011. We already miss the sound of his footsteps along the corridor, physical expression of the intensity of his thoughts. It was an honour to know him.

## FOOTNOTES

The abbreviations used are: BG, Birbeck granule; BSA, bovine serum albumin; CD, circular dichroism; CMS, cytomembrane sandwiching structures; CRD, carbohydrate-recognition domain; DCs, dendritic cells; DC-SIGN, dendritic cell-specific ICAM-3-grabbing non-integrin;  $D_{max}$ , maximal diameter; ECD, extracellular domain; EDC/NHS, N-Hydroxysuccinimide / ethyl (dimethylaminopropyl) carbodiimide; EDTA, ethylene diamine tetra-acetic acid; EGS, ethylene glycol bis(succinimidylsuccinate); FACS, fluorescence-activated cell sorting; gp120, surface HIV glycoprotein of 120 kDa ; HIV, human immunodeficiency virus; LCs, Langerhans cells; Lg, langerin; MALDI-MS, Matrix Assisted Laser Desorption Ionization Mass Spectrometry; MALDI-TOF, Matrix Assisted Laser Desorption Ionization Time Of Flight; MES, 2-(N-morpholino)ethanesulfonic acid; PEG, poly ethylene glycol; pdb, Protein Data Bank;  $R_g$ , radius of gyration; RU, resonance unit; SAXS: Small-Angle X-ray Scattering; SE-HPLC, size-exclusion high-performance liquid chromatography; SPR, surface plasmon resonance; TFA, Tri Fluoro Acetic acid; TLRs, Toll-like receptors.

---

## Chapitre 12

# Etude de l'interaction de la Langérine avec des sucres de type glycosaminoglycanes

Cette étude de l'interaction de la Langérine avec les glycosaminoglycanes se base sur la capacité de la Langérine à interagir avec des structures saccharidiques sulfatés (Figure 5.6) [99, 38]. Parmi ces structures saccharidiques particulières et les observations faites de l'interaction de la Langérine avec le keratane sulfate [99], sucre de la famille des glycosaminoglycanes (Figure 6.3), nous avons voulu mieux caractériser ce type d'interaction avec cette famille. De plus, les glycosaminoglycanes sont abondants dans le corps humain et en particulier dans la matrice extracellulaire dans laquelle la Langérine est exposée.

Les glycosaminoglycanes sont pour la plupart liés à une protéine pour former le protéoglycane. Ils sont composés d'un disaccharide de base répété différent pour chacun d'eux. On retrouve donc aussi bien des structures composées d'un galactosamine sulfaté que d'un glucosamine sulfaté démontré comme interagissant avec la Langérine [99, 38]. Pour ces différentes raisons, les GAGs sont de bon candidats pour interagir avec la Langérine.

De plus, à l'IBS, un groupe étudie les interactions entre les glycosaminoglycanes et des protéines, ils ont donc développé des outils pour ces études. Au sein d'une collaboration, nous avons pu profiter des compétences et des conseils de ce groupe. Ainsi dans ce travail, la caractérisation des différents glycosaminoglycanes a donc été réalisée par Romain Vivès et Emilie Vassal.

Enfin, une étude de modélisation a été réalisée en collaboration avec Alessandra Nuri et Antoine Daina de l'Université de Genève.

## 12.1 Développement des expériences de SPR

### 12.1.1 Développement des surfaces

Pour caractériser l'interaction de la Langérine avec les glycosaminoglycanes, nous avons utilisé la technique de résonance plasmonique de surface. Cette technique nécessite de greffer une des molécules à étudier sur la surface. Nous avons utilisé de l'héparine biotinylée à une extrémité de l'oligosaccharide qui permet de l'immobiliser sur une surface précédemment fonctionnalisée avec la streptavidine. On utilise les propriétés d'interaction entre la streptavidine et la biotine pour immobiliser l'héparine sur la surface. L'affinité de la streptavidine pour la biotine est de l'ordre de 10 nM qui permet une fixation forte de la molécule biotinylée sans avoir de dissociation dans le temps. La surface référence est seulement fonctionnalisée avec la streptavidine alors que la surface spécifique est fonctionnalisée avec la streptavidine puis l'héparine est immobilisée.

La fonctionnalisation de la surface par l'héparine est réalisée de façon classique par un couplage covalent entre les amines de la streptavidine et les carboxyles activés par le traitement EDC/NHS. L'immobilisation de l'héparine est faite à un débit de 5  $\mu\text{L}/\text{min}$  dans un tampon de course HBS-P commercial (GE Healthcare). Afin de limiter un maximum d'interaction non spécifique, l'héparine biotinylée à immobiliser est incubée dans une solution contenant 300 mM de NaCl. La réponse d'immobilisation de l'héparine est de 20 à 30 RU. Cette quantité est largement suffisante car l'héparine 6 kDa (ou 15 kDa) possède un grand nombre de sites potentiels pour l'interaction.

### 12.1.2 Amélioration de la régénération

Lors des premières expériences réalisées sur une surface fonctionnalisée avec de l'héparine, nous avons remarqué que l'EDTA n'était pas capable de régénérer complètement la surface (environ 50%). Donc nous avons dû changé de procédure, nous avons testé le SDS à une concentration de 0,05%, celui-ci permet une régénération complète et nous avons utilisé cette procédure pour les premières études. Mais lors de ces premières études de titration, nous avons remarqué que la réponse maximale pour une concentration donnée diminuait drastiquement en fonction du nombre de régénérations de la surface (Figure 12.1, rouge). Effectivement, au bout de 60 cycles nous observons une diminution du signal de l'ordre de 60%. Donc pour les études d'inhibition avec tous les GAGs testés, nous avons besoin d'une plus grande stabilité de la réponse maximum pour comparer les différents  $\text{IC}_{50}$ . Pour cela, nous avons testé différentes molécules et nous en avons trouvé une qui régénère parfaitement la surface et qui permet d'obtenir une plus grande stabilité de la réponse maximum en fonction du nombre de cycles. Effectivement à une concentration



de 333 mM de  $\text{MgCl}_2$ , la régénération est complète et inflige une diminution faible de la réactivité de la surface (Figure 12.1, bleu). En effet, au bout de 60 cycles, la réponse maximum a diminué de 5% avec le  $\text{MgCl}_2$  contrairement au 60% infligé par le SDS.

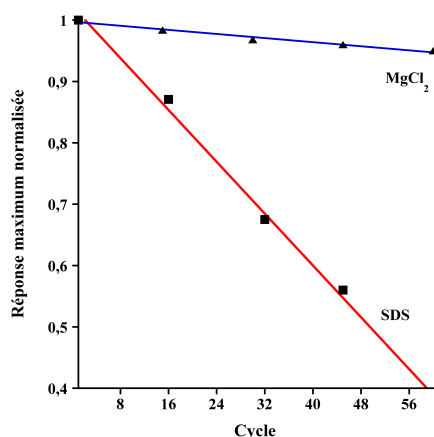


Figure 12.1 – Représentation de la réponse maximum, normalisée par rapport au RU, en fonction du nombre de cycles pour une concentration fixe de Lg-ECD (500 nM). Rouge : effet de la régénération par le SDS. Bleu : effet de la régénération par le  $\text{MgCl}_2$ .

Malheureusement cette observation n'a pas été immédiate et donc une partie des résultats présentés conserve la procédure de régénération par le SDS. Une expérience de titration nécessite une dizaine de cycles qui correspond à une perte de 10% du signal. Donc à l'échelle d'une expérience, la diminution du signal peut être acceptable.

### 12.1.3 Interaction non spécifique sur la piste témoin

Un point important avec la Langérine est la capacité de la protéine à se fixer au dextran carboxyméthylé des surfaces des sensors chip CM4. Nous décrivons cette interaction comme non spécifique mais devrait être plutôt appelée une interaction non désirée. Selon les conditions expérimentales, cette interaction non désirée avec la matrice est gênante pour l'étude, il a donc fallu déterminer des conditions de travail adaptées. Premièrement, la fonctionnalisation de la surface avec la streptavidine contribue à masquer une partie du dextran et réduit significativement cette interaction non désirée. Deuxièmement, comme vous pouvez le voir en comparant la figure 12.2 A et 12.2 B, l'interaction non spécifique est principalement dépendante du site calcium de la Langérine car en présence d'EDTA cette interaction sur le canal témoin est très faible comparée à l'interaction spécifique sur le canal immobilisé d'héparine (Figure 12.2, B).

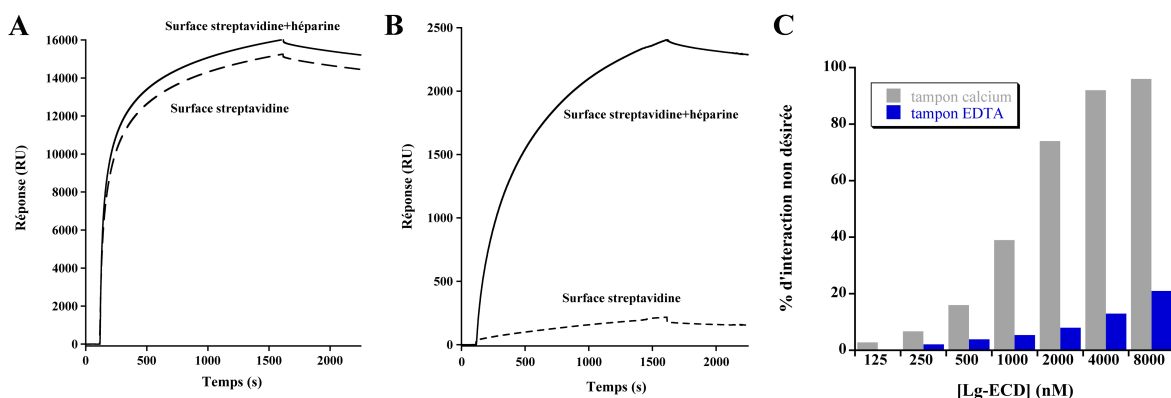


Figure 12.2 – Sensorgrammes représentant l’interaction de Lg-ECD à 2  $\mu\text{M}$  sur la surface témoin (streptavidine) et la surface d’étude (streptavidine + héparine) en tampon de course calcium (A) ou EDTA (B). C : Histogrammes des pourcentages de l’interaction non désirée par rapport au signal total pour différentes concentrations en Langérine ECD en tampon calcium (gris) et en tampon EDTA (bleu).

Cette interaction non désirée ne nous permet pas d’estimer des valeurs de  $\text{EC}_{50}$  en tampon calcium car au dessus de 2  $\mu\text{M}$  de Lg-ECD cette interaction contribue à plus de 80% du signal (Figure 12.3, C). Les expériences réalisées en présence de calcium seront conduites dans des conditions ne dépassant pas 1  $\mu\text{M}$  de Langérine ECD. Ainsi, les gammes de Langérine utilisées en tampon calcium se placent dans une échelle de concentration où l’interaction non désirée sera inférieure à 50% du signal total (Figure 12.3, C), ceci permet d’extraire, par soustraction des canaux témoin et de travail, une composante spécifique de l’interaction qui a du sens.

Par contre en tampon EDTA, une telle interaction non spécifique avec la surface dextran n’est pas ou peu présente et ne gêne pas la conduite des expériences. En effet, il faut monter jusqu’à 8  $\mu\text{M}$  de Langérine pour atteindre 20% d’interaction sur la piste témoin par rapport à la piste de travail. Ceci se prend donc parfaitement en compte pour une soustraction.

Il se trouve que la concentration de 1  $\mu\text{M}$  n’est pas suffisante pour atteindre la saturation et estimer un  $\text{EC}_{50}$ . C’est pour cela que nous avons estimé un  $\text{EC}_{50}$  de l’interaction Langérine/héparine qu’en tampon EDTA.

En fin de thèse et suite à de nombreuses discussions avec les utilisateurs de la SPR à l’institut, j’ai validé une procédure d’immobilisation de la streptavidine différente qui permet de réduire massivement cette interaction non-désirée avec la Langérine [72]. Cette procédure consiste à activer la surface par le mélange EDC/NHS puis de fixer de l’éthanolamine, cette étape est réalisée trois fois consécutivement avant de fixer la streptavidine (Figure 12.3 A). Ceci permet de réduire considérablement le nombre de fonction carboxylique présent le long des chaînes de dextran sans pour autant réduire la capacité de

fixation de la streptavidine.

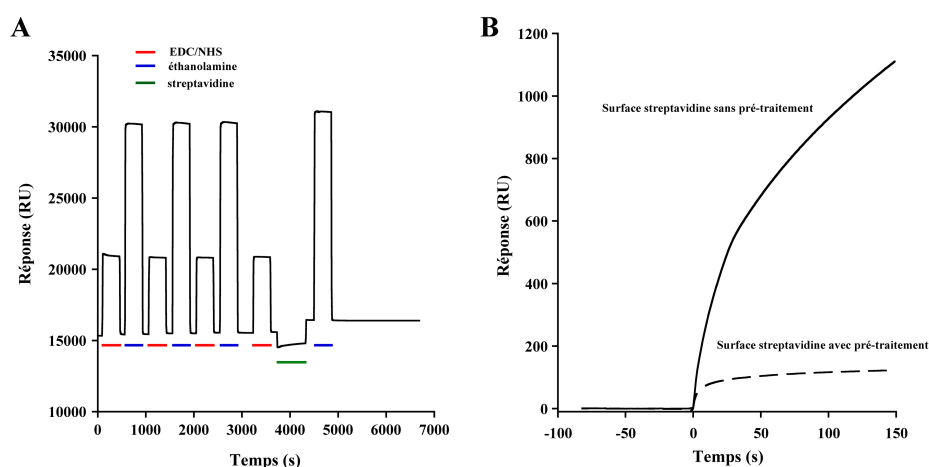


Figure 12.3 – Sensorgrammes de la greffe de la streptavidine précédée du pré-traitement (A) et de l'interaction non-désirée de la Langérine à une concentration de 500 nM avec ou sans le pré-traitement (B).

Il est clair que ce pré-traitement permet de diminuer significativement cette interaction non-désirée de l'ordre de 90% (Figure 12.3 B). Donc dans l'avenir des études de SPR avec la Langérine, il me semble indispensable d'utiliser cette procédure afin de pouvoir utiliser des concentrations supérieures à 1  $\mu$ M en tampon calcium et donc espérer saturer la surface d'étude pour déterminer des  $EC_{50}$ .

## 12.2 Un nouveau mode d'interaction de la Langérine avec les glycosaminoglycanes

Cette partie, sous forme d'un article, présente les résultats obtenus dans l'étude de l'interaction de la Langérine avec les glycosaminoglycanes.

Nous avons donc peu mettre en évidence cette interaction mais surtout les particularités qui y sont associées. En effet, nous avons montré que la Langérine reconnaît les glycosaminoglycanes mais avec un mode d'interaction original jamais décrit jusqu'à lors :

1. L'interaction ne s'effectue pas par le biais du site calcique.
2. Cette interaction nécessite la structure oligomérique de la Langérine.
3. La Langérine est capable de reconnaître de préférences les structures de type héparine puis chondroïtine.
4. Le degré de sulfate des glycosaminoglycanes influe significativement sur l'affinité.
5. Le sulfate en position C6 des glucosamines jouent un rôle important dans le mode d'interaction mais sans être indispensable.

Ce travail expérimental est complété par une proposition de site d'interaction suggéré par modélisation moléculaire, réalisée en collaboration avec Alessandra Nurisso et Antoine Diana. Ce dernier est en parfait accord avec les données biochimiques.

# **A NEW $\text{Ca}^{2+}$ -INDEPENDENT INTERACTING MODE OF LANGERIN WITH GLYCOSAMINOGLYCANS CONTRASTS WITH gp120 BINDING MODE.**

Eric Chabrol<sup>a,c,d</sup>, Alessandra Nurisso<sup>e</sup>, Antoine Daina<sup>e,h</sup>, Emilie Vassal-  
Stermann<sup>b,c,d</sup>, Michel Thepaut<sup>a,c,d</sup>, Eric Girard<sup>c,d,f</sup>, Romain R. Vivès<sup>b,c,d</sup>, Franck  
Fieschi<sup>a,c,d,g\*</sup>

<sup>a</sup>Groupe Membrane & Pathogens, <sup>b</sup>Groupe SAGAG, CEA, DSV and <sup>f</sup>Groupe ELMA, Institut  
de Biologie Structurale, 41 rue Jules Horowitz, Grenoble, F-38027, France; <sup>c</sup>CNRS, UMR  
5075, Grenoble, France; <sup>d</sup>Université de Grenoble, Grenoble, France, <sup>e</sup>Université de Genève,  
Département de pharmacochimie, 30 quai Ernest-Ansermet, Genève, CH-1211, Switzerland,  
<sup>g</sup>Institut Universitaire de France, 103 boulevard Saint-Michel, 75005 Paris, France,  
<sup>h</sup>Molecular Modeling Group, Swiss Institute of Bioinformatics, Quartier Sorge-Batiment  
Genopode, CH-1015 Lausanne, Switzerland.

Running head: New specificity and binding mode of langerin towards GAGs.

\*Address correspondence to: Franck Fieschi, Institut de Biologie Structurale (IBS), 41 rue  
Jules Horowitz, Grenoble, F-38027, France. Tel: +334 3878 91 77. Fax: +33 4 38 78 54 94.  
E-mail: [franck.fieschi@ibs.fr](mailto:franck.fieschi@ibs.fr)

## **Abstract**

Langerin is a C-type lectin specifically expressed in Langerhans cells. As recently shown for HIV, langerin is thought to capture pathogens and mediate their internalisation into Birbeck Granules for elimination. However, the precise functions of langerin remain elusive, mostly because of the lack of information on its binding properties and physiological ligands. Based on recent reports that langerin binds to sulfated sugars, we conducted here a comparative analysis of langerin interaction with mannose-rich HIV glycoprotein gp120 and glycosaminoglycan (GAGs), a family of sulfated polysaccharides expressed at the surface of most mammalian cells. Our results first revealed that langerin bound to these different glycans through very distinct mechanisms and led to the identification of a novel, GAG-specific binding mode within langerin. In contrast to the canonical lectin domain, this new binding site showed no  $\text{Ca}^{2+}$ -dependency, and could only be detected in entire, trimeric extracellular domains of langerin. Interestingly binding to GAGs, did not simply rely on a net charge effect, but rather on more discrete saccharide features, such as 6-O-sulfation, or iduronic acid content. Using molecular modelling simulations, we proposed a model of langerin/heparin complex, which located the GAG binding site at the interface of two of the three Carbohydrate-recognition domains of the protein, at the edge of the  $\alpha$ -helix coiled-coil. To our knowledge, the binding properties that we have highlighted here for langerin, have never been reported for C-type lectins before. These findings provide new insights towards the understanding of langerin biological functions.

**Keywords :** C-type Lectin; Dendritic cells; HIV; Molecular modeling; SPR, Heparin.

## INTRODUCTION

Langerin is a C-type lectin receptor highly expressed in Langerhans cells (LCs), a subset of dendritic cells, which reside in skin epidermis and mucosal epithelium. From the N- to the C-terminus, langerin is composed of a short cytoplasmic region, a unique transmembrane domain and a large extracellular domain (ECD) subdivided into a neck domain and a C-terminal carbohydrate-recognition domain (CRD). Initially identified as a molecular marker of LCs (recognized by the LC-specific DCGM4 monoclonal Antibody) <sup>1</sup>, langerin initially caught attention, a decade ago, for its unique ability to promote, by itself, the formation of a specific organelle, only present in LCs, the Birbeck Granule <sup>2</sup>. More recently, this feature was further highlighted by the observation that langerin was able to prevent HIV transmission to T-cells following direct interaction with gp120 and internalization of the virus within Birbeck Granule for elimination <sup>3</sup>. The implication of langerin in the prevention of HIV transmission strongly contrasts with the fate of HIV particles interacting with DC-SIGN, another C-type lectin receptor of the same family <sup>4</sup>. Indeed, DC-SIGN, which is present at the surface of another subtype of dendritic cells, is largely described as an important factor promoting trans-infection of HIV particles from DCs to T-cells <sup>5,6</sup> and is therefore considered as critical in the initial steps of HIV transmission. langerin-expressing Langerhans cells are present in epidermis, the upper layer of skin and mucosa and are therefore the first cell subsets encountering the virus while DC-SIGN, expressed in immature interstitial DCs, is present in dermis and in the deeper layer of mucosa <sup>7</sup>. DC-SIGN has become a target for potential microbicides for many chemical consortiums which intend to develop inhibitors of the initial step of HIV transmission <sup>8,9-14</sup>. However, it seems that, besides being a powerful DC-SIGN inhibitor, the perfect compound should also have no effect on langerin function to preserve the efficacy of the natural mucosal barrier towards HIV genital infection. A research consortium to which we belong has been jointly working along these lines with some preliminary successes <sup>13,15</sup>.

To support these developments, but also to better understand the biological role of langerin, a good knowledge of its binding properties together with the identification of natural and potentially physiological ligands is essential. Some glycan arrays studies have already shed light on the specific properties of langerin glycan recognition <sup>16,17,18</sup>.

Within its  $\text{Ca}^{2+}$  binding site, langerin is able to recognize oligosaccharides with terminal Man or GlcNAc and also, with some restriction, oligosaccharides involving terminal

fucose linked to galactose<sup>16,18</sup>. DC-SIGN can bind to internal Man within large oligosaccharides and can also recognize a wide variety of fucose from Lewis antigen derivatives<sup>19</sup>. Indeed, langerin and DC-SIGN share the ability to bind high mannose structures, as found on the HIV gp120 envelope protein.

T. Feizi, the Nobel laureate R. Steinman *et al.* have been the first to highlight the specificity of the murine langerin towards sulfated sugars<sup>20</sup>. This study describes the specificity towards Lewis x analogs, which harbor terminal galactoses sulfated at position 6, whereas no recognition was observed with sugars sulfated at position 3 (in contrast with selectin). Moreover, sulfated dextrans also bind langerin, suggesting that sulfated glucose polymers could be ligands as well. Tateno *et al.* recently confirmed such a specific recognition of 6-sulfated galactose by langerin<sup>21</sup> and explored the binding properties of langerin for keratan sulfate, which is a naturally Gal-6-sulfated polysaccharide of the glycosaminoglycan (GAG) family. Indeed, apart from HIV, this study proposed keratan sulfate as a potential physiological binding partner for langerin.

From these preliminary observations, we postulated that other sulfated GAGs (such as heparan sulfate and chondroitin sulfate) might be potential ligands for langerin. In this work, we analysed the binding properties of langerin towards gp120 HIV envelope protein as well as the recognition of a large set of GAGs. From these comparative analyses, we firstly demonstrated that langerin is able to bind to a broad range of GAGs with a marked preference for heparin and heparan sulfate. More surprisingly, we observed that gp120 and GAG recognition are based on totally different modes of interactions. Through its high mannose glycans, gp120 directly interacts with the  $\text{Ca}^{2+}$  dependent-binding site of the carbohydrate domain, whereas binding to GAGs appears to be totally  $\text{Ca}^{2+}$  independent. In addition, while langerin isolated monomeric CRD is able to recognize gp120, the trimeric form of the entire extracellular domain, including CRD and neck domain, is required for the interaction with GAGs. Molecular modeling simulation was used to identify putative GAG binding sites within the protein and highlighted a most probable binding site, which location was in agreement with  $\text{Ca}^{2+}$  independency and oligomeric requirements for intermolecular recognition.

To our knowledge, such structural requirements for sugar binding have never been reported for a C-type lectin receptor (CLR). As such, our study describes a novel binding mode, and addresses new questions regarding the physiology of LCs in epidermis and mucosa, which contain large amount of GAGs.



## EXPERIMENTAL PROCEDURES

### *Expression and purification of recombinant langerin domains*

Soluble Lg-S-CRD was expressed in the periplasmic compartment and purified as previously described, using a one step Strep-Tag II purification<sup>22</sup>. Langerin ECD was expressed in inclusion bodies. Refolding and purification procedures were performed as already described<sup>23</sup>. After refolding, purification of functional Lg-ECD proteins was achieved by affinity chromatography on a mannan-agarose column (Sigma) equilibrated in buffer A (150 mM NaCl; 25 mM Tris pH 7.8) supplemented with 4 mM CaCl<sub>2</sub> and eluted in buffer A without CaCl<sub>2</sub> but supplemented with 10 mM EDTA. This step was followed by a Superose 6 size exclusion chromatography equilibrated in buffer A with 4 mM CaCl<sub>2</sub>.

### *Surface plasmon resonance*

*Interaction between langerin and gp120* - All experiments were performed on a BIAcore 3000 using CM4 chips and the corresponding reagents from BIAcore. Three different sensor chips were used to study the interaction between langerin and gp120<sub>YU2</sub>: the first one was used to study the interaction between Lg-S-CRD and gp120<sub>YU2</sub>. For this, two flow cells of a CM4 sensor chip were activated by 50 µL of a mixture EDC/NHS. Flow cell one was functionalized with 10 µg/mL BSA, in 5 mM sodium acetate pH 4.5 buffer (639 RU immobilized), blocked with 50 µL of 1 M ethanolamine and used as a control surface. Flow cell two was functionalized with gp120<sub>YU2</sub> at 10 µg/mL in 5 mM sodium acetate pH 4.5 buffer to reach an immobilization level of 514 RU. Running buffer was buffer A supplemented with 4 mM CaCl<sub>2</sub> and 0.005% of P20 surfactant. Binding assays were performed at 20 µl/min, by injecting two-fold dilution series of Lg-S-CRD (12.5 µM to 400 µM) and Lg-ECD (2 nM to 2 µM) for 2.5 and 12.5 minutes, respectively. Regeneration was achieved by injection of 50 mM EDTA (20 and 50 µl for Lg-S-CRD and Lg-ECD, respectively). The 1:1 Langmuir model, included within the BIAeval 3.1 software, was used to fit the Lg-S-CRDs sensorgrams. Analysis of Lg ECD onto a gp120-immobilized surface was performed similarly. BSA negative control and gp120 surfaces (600 RU and 838 RU immobilized, respectively) were prepared as described above. Two-fold dilution series of Lg-ECD (0,48 nM to 2 µM) in running buffer A plus 4 mM CaCl<sub>2</sub> and 0.005% P20 were injected

over both surfaces at 20  $\mu\text{L}/\text{min}$  for 12.5 minutes (750  $\mu\text{L}$ ). Regeneration was achieved by injection of 50 mM EDTA for 2.5 minutes.

*Interaction between langerin and glycosaminoglycans* - Binding of langerin to GAGs was analysed using direct interaction and competition approaches. For this, three flow cells were activated as described above, and functionalized with streptavidin at 100  $\mu\text{g}/\text{mL}$  in 5 mM sodium acetate pH 4.5 buffer. Flow cell one was used as negative control surface. Fifteen kDa and 6 kDa heparins (Sigma-Aldrich) were biotinylated as described previously<sup>24</sup>, then immobilized on flow cells 2 and 3 (20 and 30 RU, respectively), by injection at 5  $\mu\text{g}/\text{mL}$  in 0.3 M NaCl for 5 minutes. Non-specific binding was removed by injection of 20  $\mu\text{L}$  of 2 M NaCl. Binding assays were performed at 10  $\mu\text{L}/\text{min}$  in running buffer A plus 4 mM  $\text{CaCl}_2$ , 0.005% P20 or running buffer A plus 1 mM EDTA, 0.005% P20. Regeneration was performed by injection of 350 mM  $\text{MgCl}_2$  for 5 minutes. For direct interaction studies, 2-fold dilution series of Lg-S-CRD (1.6 to 100  $\mu\text{M}$ ) were injected over the surfaces at a 10  $\mu\text{L}/\text{min}$  flow rate. For direct interaction studies with Lg-ECD, concentration ranges used were from 0.5 nM to 1  $\mu\text{M}$  in  $\text{Ca}^{2+}$  buffer and from 0.5 nM to 8  $\mu\text{M}$  in EDTA buffer with 2-fold serial dilution factor. For competition assays, two flow cells of a CM4 sensor chips were functionalized with streptavidin as described above, and 20-30 RU of biotinylated 6 kDa heparin were immobilized onto flow cell 2. The protein concentration used in the injected analyte sample was 500 nM together with different GAGs concentrations from 7.8 nM to 2  $\mu\text{M}$  with 2-fold serial dilution factor.

#### *Preparation of 6-O-desulfated heparin and disaccharide analysis of GAGs*

Vectors pcDNA3.1/Myc-His(-) encoding for HSulf-2 were used to transfect FreeStyle 293-F cells (Invitrogen), using the protocol provided by the manufacturer. 72 h post-transfection, selection of stable transfectants was carried out for 3 weeks by addition of G418 (400  $\mu\text{g}/\text{ml}$ ). Transfected cell culture supernatants were collected, extensively dialysed against 50 mM Tris pH 7.5, and concentrated 100X by ultrafiltration. HSulf-2 activity was assessed as previously described<sup>25</sup>, by incubation of HSulf-2 with 10 mM of fluorogenic pseudosubstrate 4-MUS (Sigma) in 50 mM Tris, 20 mM  $\text{MgCl}_2$  pH 7.5 for 3 hours at 37°C, and measurement of umbelliferone fluorescence (exc. 360 nm, em. 460 nm). Specific 6-O-desulfation of 15 kDa heparin (Sigma) was then performed by incubation with 150  $\mu\text{L}$  of

HSulf-2 supplemented with  $\text{MgCl}_2$  (2 mM final concentration) for 24 hours at 37°C. Samples were then heated at 100°C for 5 minutes to terminate the reaction.

Disaccharide analysis of heparin was performed as previously described<sup>26</sup>. Briefly, samples were exhaustively digested by successive incubation with heparinase I (10 mIU 24 h at 30°C), then heparinases II and III (10 mIU each, 24 h at 37°C), in 100 mM sodium acetate, 0.5 mM calcium acetate, pH 7.1. Complete digestion into disaccharides was confirmed by analysis of the digestion products in size exclusion chromatography, using twinned Superdex Peptide 10/300GL columns equilibrated in PBS, 0.3 M NaCl and run at 0.5 ml/min. Disaccharides were then resolved by strong anion exchange (SAX)-HPLC (Propac PA1, Dionex) equilibrated in  $\text{H}_2\text{O}$  pH 3.5, over a 0-1 M NaCl gradient. Peaks were detected by measuring absorbance at 232 nm and elution positions compared to those of authentic disaccharide standards (Iduron).

Disaccharide analysis of CS-A, DS and CS-C (Sigma) was performed similarly. Sample digestion was achieved by incubation in 50 mM Tris-HCl, 50 mM NaCl, 2 mM  $\text{CaCl}_2$ , 0.01% BSA, pH 7.5 with 500 mU of chondroitinase ABC for 24 h at 37°C. Identification of disaccharides by SAX-HPLC was carried out using a 0-0.75 M NaCl gradient calibrated with CS/DS disaccharide standards (Iduron).

### *Molecular modeling*

*Modeling of the langerin ECD target protein* - The langerin ECD structure was built by merging the crystal structure of the langerin trimer taken from the Brookhaven Protein Data Bank (PDB ID: 3kqg) with the trimeric coiled-coil neck region derived from our previously published homology model<sup>16,23</sup>. Hydrogen atoms were added and Gasteiger-Huckel partial atomic charges were assigned. Titratable groups were considered in their standard protonation state at neutral pH. Orientation of hydrogen atoms were optimized by using the AMBER force field as implemented in the SYBYL8 package (Tripos Inc., St Louis, MO).

*Surface mapping and EADock runs: identification of potential heparin interacting areas on the langerin ECD target protein* - Connolly surface was computed by the MOLCAD program<sup>27</sup> around the modeled langerin ECD structure. Electrostatic potential was mapped and visualized within the SYBYL graphical environment. The EADock DSS docking

methodology was used through the SwissDock web service <sup>28,29</sup> (<http://www.swissdock.ch>) to predict the most favorable anchoring positions for the negatively charged moieties of heparin at the langerin surface. A methylsulfate input fragment was considered herein as a probe for blind docking simulations in the *accurate* mode setting. The protein was considered as rigid. Since this simulation aimed mainly at considering the various areas important for intermolecular electrostatic recognition forces, assigned charges were carefully checked. Indeed the MMFF94 force field <sup>30</sup> delocalizes charges on the sulfate group correctly (-0.8167 on each oxygen atoms and +1.6337 on the sulfur) and the CHARMM22 force field <sup>31</sup> treats the calcium ions of langerin properly by assigning a +2.0 charge per atom. Four runs, each one yielding 250 docking solutions, were performed and docking poses were clustered for evaluation at 2Å-rmsd.

*Heparin fragments: constructions and parameterization* - XYZ coordinates of 2-N-sulfated, 6-O-sulfated  $\alpha$ -D-Glc and 2-O-sulfated  $\beta$ -L-IdoA monomers were retrieved from the Monosaccharide databank available at [www.cermav.cnrs.fr/cgi-bin/monos/monos.cgi](http://www.cermav.cnrs.fr/cgi-bin/monos/monos.cgi). Heparin fragments were then generated by alternating such monomers to form heparin tetramers, also considering the occurrence of <sup>2</sup>S<sub>O</sub> and <sup>1</sup>C<sub>4</sub> ring shapes for IdoA (Fig. 4). Partial charges were assigned according to the PIM force field <sup>32</sup> whereas geometries were adjusted through conjugate gradient energy minimization within the Tripos force field <sup>33</sup>.

*Heparin fragment and the langerin ECD target protein: molecular docking* – Heparin fragments were considered as ligands for docking performed by the AutoDock 3.0 program <sup>34</sup>. This methodology is widely considered suited for carbohydrate-protein molecular recognition and, in particular, for those involving glycosaminoglycans <sup>35</sup>. Langerin target structure and heparin fragments were herein modeled with explicit hydrogen atoms, protonated as at pH 7.4. Gasteiger-Huckel and PIM partial atomic charges were assigned to the protein and to the ligands, respectively. The three calcium atoms located at the edge of langerin were considered as separated bodies by assigning the formal charge +2 and Lennard-Jones parameters as described in our previous work <sup>23</sup>. The most favored area where heparin interactions may take place, identified through the surface mapping analysis and EADock DSS runs, was considered as potential binding site, thus enclosed in 0.4 Å-lattice grids (160Åx100Åx100Å).

During Autodock runs the protein was considered as a rigid body whereas all the rotatable bonds of the fragments were flexible, allowing the glycosidic linkages of the heparin units to adopt different yet realistic conformations. For each fragment, 50 docking solutions were

retrieved after Lamarckian Genetic Algorithm cycles, each one characterized by a generation of 50 individuals and  $1 \times 10^6$  energetic evaluations. Since water was not modeled explicitly, a dielectric constant of 4r was accounted. Finally, 2.5Å-rmsd clusters of docking solutions were established and analyzed.

*Modeling of the heparin decamer in complex with langerin ECD target protein* - An extended heparin chain was obtained by merging representative saccharidic fragment docking solutions from Autodock population clusters. The heparin chain, consisting of 10 monosaccharides, was modeled according to the energetically accessible glycosidic linkage information available in literature <sup>36</sup>. Heparin-langerin reformed complex was then submitted to staged-energy minimization cycles of 1000 iterations each, by first involving hydrogen atoms and side chain residues of langerin. Then the entire complex was free to relax. All the minimization cycles were performed within the Tripos force field added with PIM parameters, dedicated to carbohydrates <sup>32</sup>. The permittivity was set as a distance-dependant function and a Powell-type minimizer was used through the calculations.

## RESULTS

### *Analysis of langerin/heparin interaction by SPR.*

Langerin is a lectin able to recognize various sulfated carbohydrates. Among natural carbohydrate polymers harboring sulfate groups, GAGs are widely distributed in epithelia where Langerhans cells are present <sup>37</sup> and could thus candidate for langerin natural ligands. To investigate this, we first analysed binding of langerin to heparin by SPR. On a CM4 sensor chip functionalised with saturating amounts of streptavidin, we immobilized 6 kDa (Fig.1) and 15 kDa heparin (not shown) onto different flow cells up to 30 RU. langerin ECD interacted with both heparin surfaces but surface regeneration with EDTA was not complete, suggesting that the interaction between langerin and heparin is more complex than classical C-type lectin carbohydrate interactions, which are often strictly  $\text{Ca}^{2+}$  dependent. However, complete regeneration was finally obtained with a flash injection of  $\text{MgCl}_2$ . To characterize further the binding properties of langerin to heparin, we conducted a parallel comparison with the langerin/gp120 interaction, which was expected to be  $\text{Ca}^{2+}$  dependent as for classical C-type lectin-glycoconjugate complex.

### *Comparison of langerin interaction properties towards HIV-1 gp120 and heparin.*

Comparison of langerin interaction for heparin and gp120 clearly states that both ligands interact through different binding modes. Langerin interaction with the gp120 surface displays the canonical behavior awaited for C-type lectin receptors (CLRs) towards classical glycoconjugates. Indeed, CRD is able to bind to gp120 and represents the domain unit of recognition binding (Fig. 2A) and a clear affinity improvement is observed, through an avidity effect, using the whole trimeric extracellular domain (Fig. 2B). Due to the complexity of the Lg ECD multivalent interaction, no  $K_d$  could be determined. However, from the range of concentrations tested, an  $EC_{50}$  was estimated around  $282 \pm 3$  nM for the interaction with the gp120 surface (Fig. 2C). In contrast, sensorgrams obtained for Lg-S-CRD binding to the gp120 surface could be fitted to a 1:1 Langmuir binding model and yielded a  $K_d$  of 25  $\mu$ M. These data suggest that the avidity effects triggers an increase in affinity of roughly a hundred-time. Finally, as expected, the interaction is totally calcium dependent. This was confirmed by injecting Lg ECD with either 4 mM  $Ca^{2+}$  or 1 mM EDTA in the running buffer (figure 2C). Subsequently, EDTA was used to regenerate the surface between each protein injection, for analysis of both langerin CRD and ECD (Fig. 2A and 2B).

On the contrary, the langerin/heparin interaction obeys to unconventional properties that, to our knowledge, have never been reported for CLRs. Firstly, no interaction is observed using the CRD domain while strong binding to heparin is observed with Lg ECD (Fig. 2E and 2F). This suggests that the CRD is not the minimal binding unit for heparin but that the interaction site is created only upon formation of the langerin oligomer. Secondly, comparison of the interaction in presence of  $Ca^{2+}$  or EDTA in the running buffer shows that the Lg ECD/heparin complex occurs, surprisingly, in both conditions (Fig. 2F and 2H). From both series of experiments, we generated a titration curve that highlights strong differences in binding properties depending on the conditions used. In presence of EDTA, we were able to perform a complete titration and to reach the saturation with an  $EC_{50}$  of  $150 \pm 26$  nM. In contrast, no saturation of the binding could be achieved in presence of  $Ca^{2+}$ . This suggests a more complex binding mechanism that takes place in presence of calcium, where additional events participating to the Lg ECD/heparin interaction most likely occur (Fig. 2G). Therefore, we decided to focus on the specific Lg ECD/heparin interaction observed in presence of EDTA. On its own, this  $Ca^{2+}$ -independent binding represents a new type of interaction never reported for langerin.

#### *Characterization of langerin interaction with Glycosaminoglycans*

GAGs are a family of complex polysaccharides characterised by a repeating disaccharide unit comprising a N-substituted hexosamine and an uronic acid. According to the nature of the amino sugar, 2 main subfamilies can be defined. Galactosaminoglycans include galactosamine-containing Chondroitin Sulfate (CS) and Dermatan Sulfate (DS) that can be distinguished by the nature of their uronate: either exclusively glucuronic acid (GlcA) for CS or GlcA and a proportion of its C-5 epimer iduronic acid (IdoA) for DS. For glucosaminoglycans, the amino-sugar is a glucosamine that can either be associated to a GlcA (Hyaluronic Acid, HA), or a mix of GlcA/IdoA (Heparan Sulfate, HS and Heparin, Hp), or a galactose (Gal) residue (Keratan Sulfate, KS). With the exception of HA, GAG disaccharide units can be further modified by addition of O-sulfate groups: at C-6 of Gal for KS, at C-2 of IdoA, C-6 of GlcNAc/GlcNS and occasionally C-3 of GlcNS for HS/Hp and at C-4/C-6 of GalNAc and C-2 of IdoA for CS/DS. According to sulfation patterns, CS has been sub-categorized into CS-A (preferentially 4-O-sulfated) and CS-C (preferentially 6-O-sulfated) (See Figure 3A for structure description of various GAGs).

In this study, we have assessed the ability of Lg-EC to bind to various GAGs in order to identify specific saccharide features required for the interaction. Binding properties of Lg-EC towards other GAGs was assessed by competition assays. For this, Lg-EC was pre-incubated with concentration series of free GAGs, prior to injection onto the heparin functionalised surface (Figure 3B). Maximal responses (RU<sub>max</sub>) obtained from the sensorgrams (for competition with free Hp, see figure 3B) were then used to calculate IC<sub>50</sub>s for each tested GAGs (figure 3C). Results first indicated that langerin preferentially binds to HS-type GAGs, as free Hp and HS were found to be the most potent inhibitors, with IC<sub>50</sub>s of  $30.45 \pm 13.5$  nM and  $141.5 \pm 20.5$  nM, respectively. CS/DS-type GAGs showed some inhibition, although to a much lower level. Interestingly, great discrepancies could be observed between these samples. CS-C was found to be the best inhibitor (IC<sub>50</sub> of  $423 \pm 76$  nM), followed by DS (IC<sub>50</sub>  $1.2 \pm 0.39$   $\mu$ M), and CS-A (IC<sub>50</sub> of  $3.3 \pm 0.17$   $\mu$ M). Finally, in our hands, KS failed to inhibit Lg-EC/hp interaction in the range of concentration tested (data not shown).

We then compared these binding data to the structural information obtained on these GAGs by disaccharide analysis (Table I). Interestingly, all CS/DS samples showed very similar levels of overall sulfation, indicating that binding activity could not simply be related to a net charge effect. Although CS-A and CS-C are very closely structurally related, the latter was found to be 10 times more potent an inhibitor of Lg-EC/hp interaction than CS-A. Disaccharide analysis of these 2 samples revealed that CS-C showed a greater content in 6-O

sulfation and 2-O-sulfation, but was the least 4-O-sulfated. The nature of the uronate could not be determined from heparinase-based disaccharide analysis. However, CS-C higher level of 2-O-sulfation also suggested a greater proportion of IdoA in CS-C than in CS-A. These data suggest a possible contribution of IdoA, 2-O and/or 6-O sulfates in binding to Lg-EC, but not 4-O-sulfates. Surprisingly, DS inhibited Lg-EC/Hp interaction to an intermediate level, despite being the least 6-O-sulfated of all 3 samples and having a very low level of 2-O-sulfation. One likely explanation is that DS is naturally enriched in IdoA, which compensates for the lower sulfation content.

Further structure/activity information could be obtained from the competition assays performed with Hp and HS. Again, HS showed inhibitory properties fairly close to that of Hp, despite being significantly less sulfated. This supported further the importance of sulfation pattern rather than net charge for the interaction. More interestingly, we found that treatment of Hp with HSulf-2 resulted in a 186 % decrease of its inhibitory properties. HSulf-2 is an extracellular 6-O-sulfatase that specifically targets HS or Hp at the level of [IdoA(2S)-GlcNS(6S)] trisulfated disaccharides. Accordingly, disaccharide analysis data showed a ~ 80 % reduction of this disaccharide upon HSulf-2 treatment. These data highlight the importance of such disaccharide motif for Lg-EC/Hp interaction.

*Surface mapping and EADock DSS runs: identification of potential heparin interacting areas on the langerin ECD target protein.*

Connolly surface of the langerin ECD, obtained through a meticulous merging of experimental and computational structural data, was color-coded according to the molecular electrostatic potential (Fig.5A-B). Three main positively charged areas were identified. The first one refers to the cavity created by the spatial arrangement of the three Carbohydrate Recognition Domains (CRDs) characterizing langerin, rich in Lys residues (Fig.5A, blue box). Two other positively charged zones, repeated three times over the protein surface because of the symmetric nature of langerin, are formed by two CRDs. In particular, the second region consists in a groove formed by the  $\alpha 2$  helix of one CRD and  $\alpha 1$  helix of the adjacent CRD, together with the closer loops and part of the  $\alpha$ -helix coiled-coil parts of both CRDs (Fig.5B, red box). Part of the  $\alpha$ -helix coiled-coil region of langerin also accommodates the third and less extended positively charged area (Fig.5B, yellow box).

To gain more insight into the possible binding of sulfated glycosaminoglycans (GAG), a methylsulfate probe was docked towards the langerin ECD structure with the EADock DSS program, known to be suited for blind docking simulations <sup>29</sup>. Docking results, evaluated



through a CHARMM-based energy function and clustered at 2 Å-rmsd, were visualized in the same referential as the Connolly surface. The great majority of methylsulfates were predicted to be well recognized by the positively charged areas surrounded by the red box (76.1% of all solutions) (Fig.5B) whereas the little cavity depicted by the blue box accommodates few methylsulfate fragments (5.5% of all solutions) (Fig.5A). These converging results, obtained by two unrelated computational approaches, represent robust starting points for locating the most probable areas of intermolecular recognition between langerin and heparin fragments.

*Langerin and heparin recognition: flexible docking and modeling of the decameric heparin chain.*

One main and extended region was considered for Autodock3 calculations. This one includes both the blue and red boxes as shown in the Fig. 5. Its dimension allowed accounting for only one of the three identical sites created by the symmetrical nature of langerin. Heparin fragments were chosen instead of a long heparin chain to allow full ligand flexibility during calculations.

Docking solutions from the most populated clusters, characterized by reasonable torsion angles and negative (favorable) Autodock scores, were then selected, merged together and minimized. Finally, a decameric heparin chain was obtained (Fig. 6) <sup>36</sup>. Geometrical properties (ring shapes and glycosidic torsions) referred to the heparin bound state are reported in Table II.

The bound geometry of heparin decasaccharide is predicted to adopt a double S-shape conformation, which allows for electrostatic interactions and hydrogen bonds with basic and other polar residues, involving two monomers of langerin (Fig.6B). The electropositive niches of the protein accommodate well the corresponding charged chemical partners of the decamer. In particular, the presence of two sulfate groups on each GlcNS(6S) residue of heparin seems to reinforce the binding by playing a *bridging* role between both proteic chains. The GlcNS(6S) residue at the reducing end, for example, is trapped in salt bridges involving the negatively charged oxygen of the sulfate at position 6 and the side chain of Arg187 (langerin chain B, Fig.6B). The same residue is characterized by a salt bridge formed between the oxygen of the sulfate at position 2 and the Lys186 side chain (langerin chain A, Fig.6B). Similar situations were found for the GlcNS(6S) residues, labeled 2 and 8 in Fig.4.

A more detailed analysis of the reformed complex (Table III) brings to light the importance of the O2 sulfation of GlcNS(6S) residues: all of them are implicated in salt

bridges or hydrogen bonds with polar aminoacids of langerin. Indeed, one can assume that O6 is mandatory for stabilizing the complex, beside the key bridging role of such functional groups, as described previously. The corresponding charged groups of IdoA(2S) residues also favourably influences the complex formation.

## DISCUSSION

As a C-type lectin receptor of dendritic cells, langerin is considered as an adhesion or pathogen recognition receptor. However, the real function of langerin remains elusive, mainly because its natural ligands are still unidentified. The only roles ascribed to langerin are an implication in HIV binding <sup>38</sup> and an essential involvement in the formation of Birbeck granule (BG) in Langerhans cells <sup>2,23</sup>. However, the function of such granules keeps being a matter of debate, since langerin knock-out mice lacking BGs do not display any phenotypic discrepancies (no abnormal response to *Mycobacterium tuberculosis* and *Leishmania major* infection nor to chemical induced skin carcinogenesis) <sup>39</sup>.

Up to now, the only positive roles associated with langerin and BGs in the literature is the ability of the lectin to bind and internalize HIV into BGs, thereby contributing to viral clearance<sup>3</sup>. Finally, it has been shown in LCs that langerin, along with CD1a, is also involved in the induction of cellular immune responses to *Mycobacterium leprae*, through the presentation of a non peptide antigen and a possible uptake *via* BGs <sup>40</sup>. This reinforced the previous suggestions that BGs could be a non-classical antigen-processing pathway <sup>41</sup>.

Here, we have first characterized at the molecular level the interaction between langerin and HIV envelope glycoprotein gp120. Between the single CRD domain and the whole extracellular domain comprising 3 CRDs, an apparent 100-fold rise of relative binding affinity is observed. This avidity effect, observed here for langerin oligomeric form, is well known amongst CLRs. More interesting is the  $K_d$  of 25 $\mu$ M, measured for gp120 interaction with monomeric CRD. As reported for many other CLRs, CRDs usually exhibit millimolar affinity for monovalent sugars or when recognition occurs only through an oligosaccharide terminal sugar. A  $K_d$  of 25  $\mu$ M suggests that a more extended binding must occur between gp120 glycan and langerin CRD. Thus, langerin does not solely bind the terminal mannose of the high mannose present on gp120 but rather a larger oligomannose motif. Finally, the  $EC_{50}$  of 282 nM for the Lg ECD/gp120 interaction is in good agreement with the apparent  $K_d$  reported for a langerin/gp140 interaction in another recent study<sup>42</sup>.

Apart from oligomannose, langerin has been shown to have a rather unique specificity, amongst CLRs, towards sulfated sugars <sup>20</sup>. Among putative physiological ligands, keratan sulfate (KS) has been proposed <sup>21</sup>. Indeed, KS is constituted by a repetition of LacNAc motif (Gal $\beta$ 1-4GlnAc $\beta$ 1-3)<sub>n</sub> that can be either sulfated on the C6 of the galactose or on the N-Acetyl glucosamine. As shown by Tateno et al <sup>21</sup>, the affinity of langerin for KS seems to be mainly related to its sulfation content and more particularly regarding galactose C6 sulfation ([6-SO<sub>4</sub>]Gal $\beta$ 1-4GlnAc). Reported improved binding upon de-sialylation of KS also suggests recognition through the terminal [6-SO<sub>4</sub>]Gal at the langerin Ca<sup>2+</sup> binding site. However, KS does not constitute the major GAG potentially encountered by Langerhans cells in epithelium and mucosal tissues. Considering this affinity of langerin for sulfated glycans together with the Langerhans cell location and migrating properties, we decided to define langerin binding properties towards a broader range of GAGs. Emphasis was more particularly given to HS (and structurally related heparin), which is abundant in epithelium and mucosa, and directly exposed at dendritic cells surfaces where it participates to the capture of many pathogens as well as immune activation <sup>43,44</sup>.

From then, we went from surprise to surprise. First of all, langerin is able to bind heparin, HS but also several types of CS. Secondly, this binding can be independent of Ca<sup>2+</sup> as shown by interaction studies performed in the presence of EDTA. Thirdly, affinities for heparin and heparan sulfate are in the nanomolar range, ranking them as the best ligands ever described for langerin. Last, but not least, the interaction is strictly dependent upon oligomerization and absolutely not detectable with a single monomeric CRD. From all these points, langerin/GAG interaction appears to be completely different from how related CLRs traditionally recognize their ligands. The imperative requirement for langerin trimeric form suggests the existence of a unique binding site constituted by the assembly of at least 2 of 3 protomers. The nanomolar range affinity may suggest an extended binding site. Moreover, contributions of the sulfate groups through intermolecular electrostatic interactions have also been postulated. We have also studied the interaction of langerin with other GAGs, using competition approaches. Data obtained clearly showed a selectivity of the lectin for HS-like GAGs (HS and heparin), although langerin also bound to a much more modest level to CS/DS. Interestingly, we also observed binding selectivity amongst the CS/DS samples tested, langerin exhibiting the highest binding to CS-C. Comparison of these data to GAG disaccharide analysis showed that binding to langerin could not simply be attributed to a net charge effect and that specific saccharide features were most likely required. Our results

suggest that C6 sulfation as well as iduronic acid strengthen the binding. Moreover, the affinity loss observed for heparin upon HSulf-2 treatment highlights the importance of the C6 sulfate present in the [IdoA(2S)-GlcNS(6S)] motif.

We used the recent crystal structure of the langerin trimer<sup>16</sup> to undertake molecular modeling analysis of langerin interaction with heparin fragments. Combining the trimeric X-ray structure of a truncated ECD with the previously modeled neck region yielded a reasonably robust model to initiate the search for putative favorable heparin binding regions. Two main areas of interaction with heparin have been identified on the whole langerin surface through MOLCAD electrostatic potential analysis and EADock DSS cavity detection and blind methylsulfate docking. Thanks to this preliminary dual approach, the more precise Autodock docking of heparin fragments was restricted to those specific areas. Outcome results of these molecular simulations yielded three main conclusions: i) neither methylsulfate nor heparin fragment docking pose interact with the calcium ions; ii) heparin fragment-langerin interactions are driven by direct polar forces (salt bridges, hydrogen bonds); iii) the molecular recognition of heparin fragments depends upon more than one langerin CRD : the most populated docking clusters occupy both CRDs characterizing the edge of the  $\alpha$ -helix coiled-coil (Fig.5B, red box).

Building on these clear modeling output, it was then possible to construct straightforwardly a heparin decamer *in situ*. In the model, the double sulfation of GlcNS(6S) residues appeared essential for the interaction, acting also as a *bridge* between both CRDs. Globally, the proposed model of heparin/langerin molecular recognition is in full accordance with the biochemical results. However in order to get an accurate estimation of the free energy of binding and to go further towards a physically relevant description of molecular recognition, molecular dynamics studies would be considered as suitable to take into account the charged and flexible aminoacids coating the binding region. Moreover, the construction of the heparin chain was limited to ten monomers. Modeling of a more extended heparin chain could involve other areas of the protein, for instance at the top of the described region, toward the electropositive *cavity* (Fig.5 A) involving Lys 299 and Lys 313 residues<sup>21</sup>.

The multiple approaches of our work gives convergent evidence for a novel binding mode of langerin ligands. Remarkably, the binding is independent from the canonical  $\text{Ca}^{2+}$ -site. Previously, the existence of two distinct binding sites within langerin has been postulated on the basis of an X-ray structure of langerin in complex with maltose. A maltose was described onto the  $\text{Ca}^{2+}$  but also within a specific cleft present in langerin CRD only<sup>45</sup>. It finally turns out that the electron density initially attributed to maltose in this large cleft

without  $\text{Ca}^{2+}$  was in fact the C-terminus of an affinity tag coming from a neighboring molecule in the crystal lattice<sup>17</sup>. However, although this initial proposal for a second binding site for carbohydrates, independent of  $\text{Ca}^{2+}$  was not validated, it finds here, in a different area of the protein, a new revival. Its nature is totally new in CLRs since it represents, as far as we know, the first binding site generated at the interface between two protomers of C-type lectin receptors. Langerin is thus able to selectively interact with sulfated carbohydrate through two totally distinct modes: i) a  $\text{Ca}^{2+}$ -dependent binding mode in the CLR canonical site when OH groups are available in C3 and C4 of the saccharide ring (as for galactose 6 sulfate, for instance) and ii) in a  $\text{Ca}^{2+}$ -independent mode for polysulfated glycans of the GAG family where either C3 or C4 OH groups is engaged in the polysaccharide glycosidic linkage.

Prior to this work, langerin specificity has already been assessed through several glycan array studies<sup>16,17,20,46</sup>. However, langerin binding properties towards GAGs has never been evaluated, as GAGs were missing from the glycan arrays used, except for the work by Tateno *et al.* In that case, heparin -as well as HS, DS, CSA and KS- were present onto the micro array, but only KS, through binding of its terminal saccharide to the canonical binding site, was identified. This result is in apparent disagreement with our present data. However, one likely explanation is that Tateno *et al.* microarray screening was conducted with an Lg-CRD-Fc fusion protein that exhibits the canonical binding site, but not the newly identified GAG binding site described here. This latter one requires the trimeric form of the protein dependant on the presence of the neck region of the extracellular domain of langerin. This critical observation clearly demonstrates the importance of CLR oligomeric organization, which cannot simply be considered as a sum of independent CRDs. Here, langerin trimerisation of CRDs also creates a new and unrelated site thanks to the neck domain of the protein.

The identification of the langerin specificity towards GAGs raises the question of the physiological relevance and role of such an interaction. HS is abundantly present in the tissues hosting Langerhans cells. Surface of dendritic cells themselves exposes proteoglycans bearing long GAG chains. Therefore, it is most likely langerin will be in contact with GAGs during the life cycle of the Langerhans cell. Interestingly, a previous work studying the biochemistry of LC trafficking pointed out that heparin, and more particularly N-sulfated glucosamine moieties of heparin, could inhibit LC trafficking<sup>47</sup>. Indeed, a heparin binding factor was postulated to be involved in LC migration. Future work will have to examine a possible role of langerin in the modulation of LC trafficking. Another possibility might be a synergistic implication of both heparin and langerin in pathogen recognition.

This work highlighted the unique properties of langerin to interact with glycans through both a  $\text{Ca}^{2+}$  binding site, as for gp120 high mannose, and a new and never reported GAG specific binding domain. This raises many new issues about the physiological role of langerin within the Langerhans cells.

## REFERENCES

1. Valladeau, J., Duvert-Frances, V., Pin, J.J., Dezutter-Dambuyant, C., Vincent, C., Massacrier, C., Vincent, J., Yoneda, K., Banchereau, J., Caux, C., Davoust, J. & Saeland, S. (1999). The monoclonal antibody DCGM4 recognizes Langerin, a protein specific of Langerhans cells, and is rapidly internalized from the cell surface. *Eur J Immunol* **29**, 2695–2704
2. Valladeau, J., Ravel, O., Dezutter-Dambuyant, C., Moore, K., Kleijmeer, M., Liu, Y., Duvert-Frances, V., Vincent, C., Schmitt, D., Davoust, J., Caux, C., Lebecque, S. & Saeland, S. (2000). Langerin, a novel C-type lectin specific to Langerhans cells, is an endocytic receptor that induces the formation of Birbeck granules. *Immunity* **12**, 71–81
3. de Witte, L., Nabatov, A., Pion, M., Fluitsma, D., de Jong, M.A.W.P., de Gruijl, T., Piguet, V., van Kooyk, Y. & Geijtenbeek, T.B.H. (2007). Langerin is a natural barrier to HIV-1 transmission by Langerhans cells. *Nat Med* **13**, 367–371
4. Tabarani, G., Thépaut, M., Stroebel, D., Ebel, C., Vivès, C., Vachette, P., Durand, D. & Fieschi, F. (2009). DC-SIGN neck domain is a pH-sensor controlling oligomerization: SAXS and hydrodynamic studies of extracellular domain. *J Biol Chem* **284**, 21229–21240
5. Geijtenbeek, T.B., Kwon, D.S., Torensma, R., van Vliet, S.J., van Duijnhoven, G.C., Middel, J., Cornelissen, I.L., Nottet, H.S., KewalRamani, V.N., Littman, D.R., Figdor, C.G. & van Kooyk, Y. (2000). DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. *Cell* **100**, 587–597
6. Solomon, T. (2010). The multiple facets of HIV attachment to dendritic cell lectins. *Cellular Microbiology*
7. de Witte, L., Nabatov, A. & Geijtenbeek, T.B.H. (2008). Distinct roles for DC-SIGN+ dendritic cells and Langerhans cells in HIV-1 transmission. *Trends Mol Med* **14**, 12–19
8. Balzarini, J. (2007). Targeting the glycans of glycoproteins: a novel paradigm for antiviral therapy. *Nat Rev Microbiol* **5**, 583–597
9. Tabarani, G., Reina, J.J., Ebel, C., Vivès, C., Lortat-Jacob, H., Rojo, J. & Fieschi, F. (2006). Mannose hyperbranched dendritic polymers interact with clustered organization of DC-SIGN and inhibit gp120 binding. *FEBS Lett* **580**, 2402–2408
10. Reina, J.J., Sattin, S., Invernizzi, D., Mari, S., Martinez-Prats, L., Tabarani, G., Fieschi, F., Delgado, R., Nieto, P.M., Rojo, J. & Bernardi, A. (2007). 1,2-Mannobioside mimic: synthesis, DC-SIGN interaction by NMR and docking, and antiviral activity. *ChemMedChem* **2**, 1030–1036
11. Borrok, M.J. & Kiessling, L.L. (2007). Non-carbohydrate inhibitors of the lectin DC-SIGN. *J Am Chem Soc* **129**, 12780–12785
12. Becer, C.R., Gibson, M.I., Geng, J., Ilyas, R., Wallis, R., Mitchell, D.A. & Haddleton, D.M. (2010). High-affinity glycopolymer binding to human DC-SIGN and disruption of DC-SIGN interactions with HIV envelope glycoprotein. *J Am Chem Soc* **132**,

15130–15132

13. Andreini, M., Doknic, D., Sutkeviciute, I., Reina, J.J., Duan, J., Chabrol, E., Thépaut, M., Moroni, E., Doro, F., Belvisi, L., Weiser, J., Rojo, J., Fieschi, F. & Bernardi, A. (2011). Second generation of fucose-based DC-SIGN ligands : affinity improvement and specificity versus Langerin. *Org Biomol Chem* **9**, 5778–5786
14. Martínez-Avila, O., Bedoya, L.M., Marradi, M., Clavel, C., Alcamí, J. & Penadés, S. (2009). Multivalent manno-glyconanoparticles inhibit DC-SIGN-mediated HIV-1 trans-infection of human T cells. *ChemBioChem* **10**, 1806–1809
15. Berzi, A., Reina, J.J., Ottria, R., Sutkeviciute, I., Antonazzo, P., Sánchez-Navarro, M., Chabrol, E., (null), (null), Cetin, I., Rojo, J., Fieschi, F., Bernardi, A. & Clerici, M. (2012). A glycomimetic compound inhibits DC-SIGN-mediated HIV infection in cellular and cervical explant models. *AIDS* **26**, 127–137
16. Feinberg, H., Powlesland, A.S., Taylor, M.E. & Weis, W.I. (2010). Trimeric Structure of Langerin. *Journal of Biological Chemistry* **285**, 13285–13293
17. Feinberg, H., Taylor, M.E., Razi, N., McBride, R., Knirel, Y.A., Graham, S.A., Drickamer, K. & Weis, W.I. (2011). Structural basis for langerin recognition of diverse pathogen and mammalian glycans through a single binding site. *J Mol Biol* **405**, 1027–1039
18. Holla, A. & Skerra, A. (2011). Comparative analysis reveals selective recognition of glycans by the dendritic cell receptors DC-SIGN and Langerin. *Protein engineering, design & selection : PEDS* doi:10.1093/protein/gzr016
19. Guo, Y., Feinberg, H., Conroy, E., Mitchell, D.A., Alvarez, R., Blixt, O., Taylor, M.E., Weis, W.I. & Drickamer, K. (2004). Structural basis for distinct ligand-binding and targeting properties of the receptors DC-SIGN and DC-SIGNR. *Nat Struct Mol Biol* **11**, 591–598
20. Galustian, C., Park, C.G., Chai, W., Kiso, M., Bruening, S.A., Kang, Y.-S., Steinman, R.M. & Feizi, T. (2004). High and low affinity carbohydrate ligands revealed for murine SIGN-R1 by carbohydrate array and cell binding approaches, and differing specificities for SIGN-R3 and langerin. *Int Immunol* **16**, 853–866
21. Tatenno, H., Ohnishi, K., Yabe, R., Hayatsu, N., Sato, T., Takeya, M., Narimatsu, H. & Hirabayashi, J. (2010). Dual Specificity of Langerin to Sulfated and Mannosylated Glycans via a Single C-type Carbohydrate Recognition Domain. *Journal of Biological Chemistry* **285**, 6390–6400
22. Thépaut, M., Vivès, C., Pompidor, G., Kahn, R. & Fieschi, F. (2008). Overproduction, purification and preliminary crystallographic analysis of the carbohydrate-recognition domain of human langerin. *Acta Crystallogr F Struct Biol Cryst Commun* **64**, 115–118
23. Thépaut, M., Valladeau, J., Nurisso, A., Kahn, R., Arnou, B., Vivès, C., Saeland, S., Ebel, C., Monnier, C., Dezutter-Dambuyant, C., Imberty, A. & Fieschi, F. (2009). Structural studies of langerin and Birbeck granule: a macromolecular organization model. *Biochemistry* **48**, 2684–2698
24. Vivès, R.R., Sadir, R., Imberty, A., Rencurosi, A. & Lortat-Jacob, H. (2002). A kinetics and modeling study of RANTES(9-68) binding to heparin reveals a mechanism of cooperative oligomerization. *Biochemistry* **41**, 14779–14789
25. Frese, M.-A., Milz, F., Dick, M., Lamanna, W.C. & Dierks, T. (2009). Characterization of the human sulfatase Sulfl and its high affinity heparin/heparan sulfate interaction domain. *Journal of Biological Chemistry* **284**, 28033–28044
26. Pye, D.A., Vives, R.R., Turnbull, J.E., Hyde, P. & Gallagher, J.T. (1998). Heparan sulfate oligosaccharides require 6-O-sulfation for promotion of basic fibroblast growth factor mitogenic activity. *J Biol Chem* **273**, 22936–22942
27. Waldherr-Teschner, M., Goetze, T., Heiden, W., Knoblauch, M., Vollhardt, H. &

- Brickmann, J. (1992). MOLCAD-Computer Aided Visualization and Manipulation of Models in Molecular Science. *Advances in Scientific Visualization* 58–67
28. Grosdidier, A., Zoete, V. & Michielin, O. (2011). SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic Acids Res.* **39**, W270–7
29. Grosdidier, A., Zoete, V. & Michielin, O. (2009). Blind docking of 260 protein-ligand complexes with EADock 2.0. *J Comput Chem* **30**, 2021–2030
30. Halgren, T. (1996). Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94 - Halgren - 1998 - Journal of Computational Chemistry - Wiley Online Library. *Journal of computational chemistry*
31. Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S. & Karplus, M. (1983). CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem* **4**, 187–217
32. Imberty, A., Bettler, E., Karababa, M., Mazeau, K., Petrova, P. & Pérez, S. (1999). Building sugars : The sweet part of structural biology. *Perspectives in Structural Biology* 392–409
33. Clark, M., Cramer, R.D.I. & van den Opdenbosch, N. (1989). Validation of the general purpose Tripos 5.2 force field. *J Comput Chem* **10**, 982–1012
34. Jones, G., Willett, P., Glen, R. & Leach, A. (1997). Development and validation of a genetic algorithm for flexible docking1. *Journal of molecular ...*
35. Cros, S., Petitou, M., Sizun, P., Pérez, S. & Imberty, A. (1997). Combined NMR and molecular modeling study of an iduronic acid-containing trisaccharide related to antithrombotic heparin fragments. *Bioorg. Med. Chem.* **5**, 1301–1309
36. Mikhailov, D., Linhardt, R.J. & Mayo, K.H. (1997). NMR solution conformation of heparin-derived hexasaccharide. *Biochem J* **328** ( Pt 1), 51–61
37. Bernfield, M., Kokenyesi, R. & Kato, M. (1992). Biology of the Syndecans: A Family of Transmembrane Heparan Sulfate Proteoglycans - Annual Review of Cell Biology, 8(1):365. ... review of cell biology
38. Turville, S., Cameron, P., Handley, A., Lin, G., Pöhlmann, S., Doms, R. & Cunningham, A. (2002). Diversity of receptors binding HIV on dendritic cell subsets. *Nat Immunol* **3**, 975–983
39. Kissenpfennig, A., Ait-Yahia, S., Clair-Moninot, V., Stossel, H., Badell, E., Bordat, Y., Pooley, J.L., Lang, T., Prina, E., Coste, I., Gresser, O., Renno, T., Winter, N., Milon, G., Shortman, K., Romani, N., Lebecque, S., Malissen, B., Saeland, S. & Douillard, P. (2005). Disruption of the langerin/CD207 gene abolishes Birbeck granules without a marked loss of Langerhans cell function. *Mol Cell Biol* **25**, 88–99
40. Hunger, R.E., Sieling, P.A., Ochoa, M.T., Sugaya, M., Burdick, A.E., Rea, T.H., Brennan, P.J., Belisle, J.T., Blauvelt, A., Porcelli, S.A. & Modlin, R.L. (2004). Langerhans cells utilize CD1a and langerin to efficiently present nonpeptide antigens to T cells. *J Clin Invest* **113**, 701–708
41. Valladeau, J., Caux, C., Lebecque, S. & Saeland, S. (2001). [Langerin: a new lectin specific for Langerhans cells induces the formation of Birbeck granules]. *Pathol Biol (Paris)* **49**, 454–455
42. Hijazi, K., Wang, Y., Scala, C., Jeffs, S., Longstaff, C., Stieh, D., Haggarty, B., Vanham, G., Schols, D., Balzarini, J., Jones, I.M., Hoxie, J., Shattock, R. & Kelly, C.G. (2011). DC-SIGN Increases the Affinity of HIV-1 Envelope Glycoprotein Interaction with CD4. *PLoS ONE* **6**, e28307
43. Bousarghin, L., Hubert, P., Franzen, E., Jacobs, N., Boniver, J. & Delvenne, P. (2005). Human papillomavirus 16 virus-like particles use heparan sulfates to bind dendritic cells and colocalize with langerin in Langerhans cells. *J. Gen. Virol.* **86**, 1297–1305
44. de Witte, L., Zoughlami, Y., Aengeneyndt, B., David, G., van Kooyk, Y., Gissmann, L.



- & Geijtenbeek, T.B.H. (2007). Binding of human papilloma virus L1 virus-like particles to dendritic cells is mediated through heparan sulfates and induces immune activation. *Immunobiology* **212**, 679–691
45. Chatwell, L., Holla, A., Kaufer, B.B. & Skerra, A. (2008). The carbohydrate recognition domain of Langerin reveals high structural similarity with the one of DC-SIGN but an additional, calcium-independent sugar-binding site. *Mol Immunol* **45**, 1981–1994
  46. Tateno, H., Ohnishi, K., Yabe, R., Hayatsu, N., Sato, T., Takeya, M., Narimatsu, H. & Hirabayashi, J. (2010). Dual Specificity of Langerin to Sulfated and Mannosylated Glycans via a Single C-type Carbohydrate Recognition Domain. *Journal of Biological Chemistry* **285**, 6390–6400
  47. O'Sullivan, G.M., Boswell, C.M. & Halliday, G.M. (2000). Langerhans cell migration is modulated by N-sulfated glucosamine moieties in heparin. *Exp. Dermatol.* **9**, 25–33

## FOOTNOTES

This study was supported by grants from Ensemble contre le SIDA -Sidaction. Eric Chabrol was supported by a fellowship from the French Ministry of Research and Higher Education.

The abbreviations used are: CRD, carbohydrate-recognition domain; ECD, extracellular domain; DC, dendritic cells; DC-SIGN, dendritic cell-specific ICAM-3-grabbing non-integrin; EDTA, ethylene diamine tetra-acetic acid; LC, Langerhans cells; Lg, langerin; SPR, surface plasmon resonance; S-CRD, *Strep*-tagged carbohydrate-recognition domain; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; GAG, glycosaminoglycan; Hp, heparin; HS, heparin sulfate; CSA, chondroitin sulfate A; DS, dermatan sulfate; CSC, chondroitin sulfate C; KS, keratin sulfate.

## ACKNOWLEDGMENTS

We thank Prof Steve Rosen for kindly providing the PcDNA3/Myc/His/HSulf-2 vector. We also thank Dr Anne Imberty (CERMAV-CNRS) and Prof Carrupt (School of pharmaceutical sciences, University of Geneva, University of Lausanne) for technical support.

## FIGURE LEGENDS

Figure 1: **Langerin ECD interaction onto heparin.** Surface was functionalized with heparin 6 kDa. 100  $\mu$ L langerin ECD at 500 nM are injected onto the surface in a  $\text{Ca}^{2+}$  containing running buffer. Two modes of surface regenerations are tested, 1: Injection of 30  $\mu$ L of 50 mM EDTA. 2: Injection of 50  $\mu$ L of 350 mM  $\text{MgCl}_2$ .

Figure 2: **Comparison of heparin and gp120 binding mode to langerin.** A) Overlay of sensorgram showing langerin CRD interaction onto gp120<sub>YU2</sub> functionalized surface in  $\text{Ca}^{2+}$  buffer. Langerin CRD concentration range is from 400  $\mu$ M to 12,5  $\mu$ M with 2 times serial dilution. B) Overlay of sensorgram showing langerin ECD interaction onto gp120<sub>YU2</sub> functionalised surface in calcium buffer. The concentration range of langerin ECD is from 2  $\mu$ M to 7,8 nM with 2 times serial dilution. C) SPR binding analysis of gp120 interaction as a function of langerin ECD concentration. D) SPR sensorgram of langerin ECD interaction onto gp120<sub>YU2</sub> functionalised surface in calcium and EDTA buffer. Langerin ECD injection has been performed at 500 nM concentration of protein. E) Overlay of sensorgram showing langerin CRD interaction onto biotinylated 6 kDa heparin functionalised surface in calcium buffer. Langerin CRD concentration was from 100  $\mu$ M to 1.6  $\mu$ M with 2-fold serial dilution. F) Overlay of sensorgram showing of langerin ECD interaction onto biotinylated 6 kDa heparin functionalised surface in calcium buffer. The concentration range of langerin ECD is from 1  $\mu$ M to 0,49 nM with 2 times serial dilution. G) Overlay of sensorgram showing langerin ECD on 6kDa heparin surface in calcium buffer (■) and in EDTA buffer (▲). H) SPR titration experiment of langerin ECD interaction onto biotinylated 6 kDa heparin functionalised surface in EDTA buffer. The concentration range of langerin ECD is from 8  $\mu$ M to 0,488 nM with 2 times serial dilution.

Figure 3: **Interaction properties of langerin to different glycosaminoglycans.** (A-C) Structure of glycosaminoglycan disaccharides. Disaccharide units for heparin/HS (A), CS/DS (B) or KS (C). R =  $\text{COCH}_3$  or  $\text{SO}_3\text{H}$  ; X = H or  $\text{SO}_3$ . The star indicates presence of glucuronic or iduronic acid C-5 epimers. D) SPR inhibition experiment of langerin ECD/heparin interaction by 15 kDa heparin. Langerin ECD concentration was 100 nM and the heparin concentration range was from 7.8 nM to 2  $\mu$ M with 2-fold serial dilution factor. E)  $\text{IC}_{50}$  values obtained with heparin, heparin treated with HSulf2 and HS. F) Histogram representation of CS/DS  $\text{IC}_{50}$ . SPR inhibition experiment of langerin ECD/heparin interaction by 15 kDa

heparin. Langerin ECD concentration was fixed at 100 nM and the heparin concentration range was from 7.8 nM to 2  $\mu$ M with 2-fold serial dilution factor. C) Histogram representation of GAGs IC<sub>50</sub>.

Figure 4: **Heparin fragments.** 2-N-sulfated, 6-O-sulfated  $\alpha$ -D-glucopyranoside (A) and 2-O-sulfated  $\beta$ -L-*idopyranoside* monomers in its <sup>1</sup>C<sub>4</sub> (B) and <sup>2</sup>S<sub>0</sub> (C) ring shapes were considered for building heparin fragments for docking calculations. Glycosidic linkages are also indicated, defined as  $\phi = \text{O5}_i - \text{C1}_i - \text{O1}_i - \text{C4}_j$  and  $\psi = \text{C1}_i - \text{O1}_i - \text{C4}_j - \text{C5}_j$ .

Figure 5: **Three-dimensional model of langerin ECD and potential heparin docking sites.** The human langerin ECD is represented by its Connolly surface, color-coded according to the molecular electrostatic potential (from blue for negative to red for positive areas). The most probable regions for interactions with heparin are indicated with colored boxes populated with methylsulfate docking solutions, represented in *space fill* (5A: top view; 5B: side view).

Figure 6: **Heparin decasaccharide in complex with langerin.** Predicted binding mode of the heparin decasaccharide (displayed in space fill -A- and capped stick -B-) in complex with langerin. Langerin is represented with its Connolly surface color-coded according to the electrostatic potential (from blue for negative to red for positive electrostatic areas). Amino acids mediating the main interactions with the decasaccharide are labeled in white.

**Table I**

Disaccharide	Hp	Hp de6S	CS-A	CS-B	CS-C	Disaccharide
$\Delta$ HexA - GlcNAc	5.2	4.6	11	9.5	4.3	$\Delta$ HexA - GalNAc
$\Delta$ HexA - GlcNAc,6S	4.1	3.9	46.8	78.6	24.2	$\Delta$ HexA - GalNAc,4S
$\Delta$ HexA - GlcNS	6.1	10.1	41	6	52.5	$\Delta$ HexA - GalNAc,6S
$\Delta$ HexA - GlcNS,6S	12.6	9.8	0	1.2	0	$\Delta$ HexA,2S - GalNAc,4S
$\Delta$ HexA,2S - GlcNS	9.4	57.4	0.5	4.3	2.5	$\Delta$ HexA - GalNAc,4S,6S
$\Delta$ HexA,2S - GlcNS,6S	60.7	11.7	0.6	0.4	16.5	$\Delta$ HexA,2S - GalNAc,6S
$\Delta$ HexA,2S - GlcNAc	1.9	2.5				
<i>Sulfate/dp2</i>	2.4	1.9	0.9	1	1.1	<i>Sulfate/dp2</i>
<i>N-sulfation</i>	88.8	89	0.6	1.7	16.5	<i>2-O-sulfation</i>
<i>2-O-sulfation</i>	72	71.6	47.3	84.1	26.7	<i>4-O-sulfation</i>
<i>6-O-sulfation</i>	77.4	25.4	42.1	10.7	71.5	<i>6-O-sulfation</i>

**Table I - Disaccharide analysis of GAGs :** For determination of GAG composition, heparin and CS samples were exhaustively depolymerised (with heparinases I, II, III and chondroitinase ABC, respectively), and the resulting disaccharides were resolved by SAX-HPLC, using a NaCl gradient calibrated with authentic standards.

**Table II**

		$\varphi = \text{O5}_i - \text{C1}_i - \text{O1}_i - \text{C4}_j$	$\psi = \text{C1}_i - \text{O1}_i - \text{C4}_j - \text{C5}_j$
IdoA*-GlcNAc	Torsion angle 1	-107.3	-163.4
GlcNAc-IdoA°	Torsion angle 2	135.8	-89.3
IdoA°-GlcNAc	Torsion angle 3	-78.4	175.1
GlcNAc-IdoA°	Torsion angle 4	36.5	-169.3
IdoA°-GlcNAc	Torsion angle 5	-65.6	-96.0
GlcNAc-IdoA*	Torsion angle 6	63.3	-76.4
IdoA*-GlcNAc	Torsion angle 7	-56.3	-89.4
GlcNAc-IdoA°	Torsion angle 8	66.2	-168.7
IdoA°-GlcNAc	Torsion angle 9	-75.1	-112.8

**Table II** – Details about the geometries of the heparin deca-saccharide in complex with langerin. The torsion angles were defined as follow:  $\varphi = \text{O5}_i - \text{C1}_i - \text{O1}_i - \text{C4}_j$  and  $\psi = \text{C1}_i - \text{O1}_i - \text{C4}_j - \text{C5}_j$ . The numbering for the torsion angles is reported in Fig.1. (\*)*L-Idopyranoside* monosaccharides in  $^1\text{C}_4$ . (°)*L-Idopyranoside* monosaccharides in  $^2\text{S}_0$ .

**Figure 1**

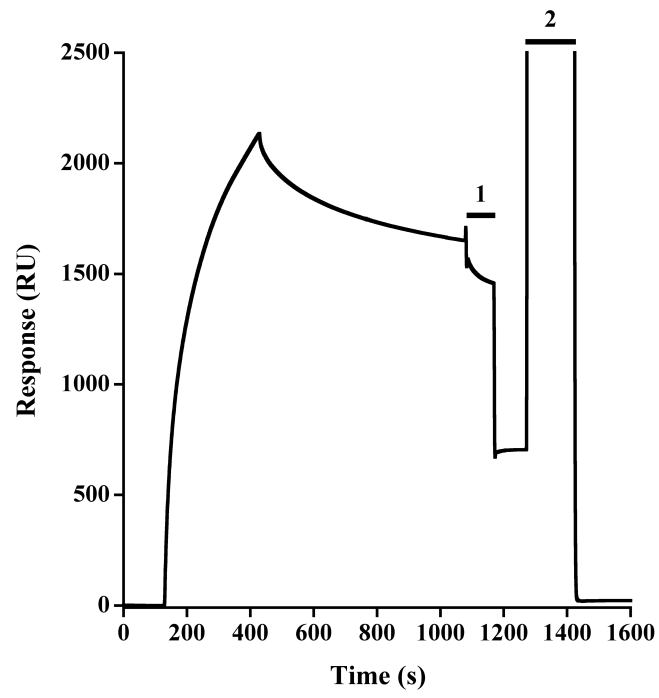


Figure 2

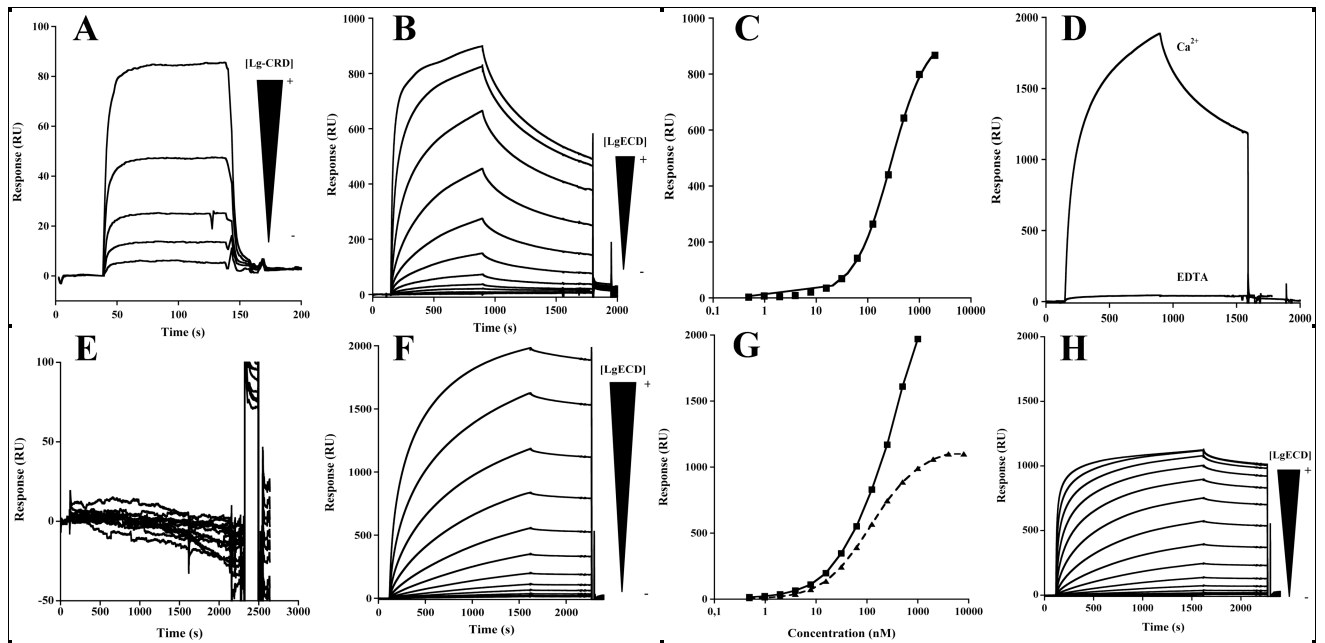
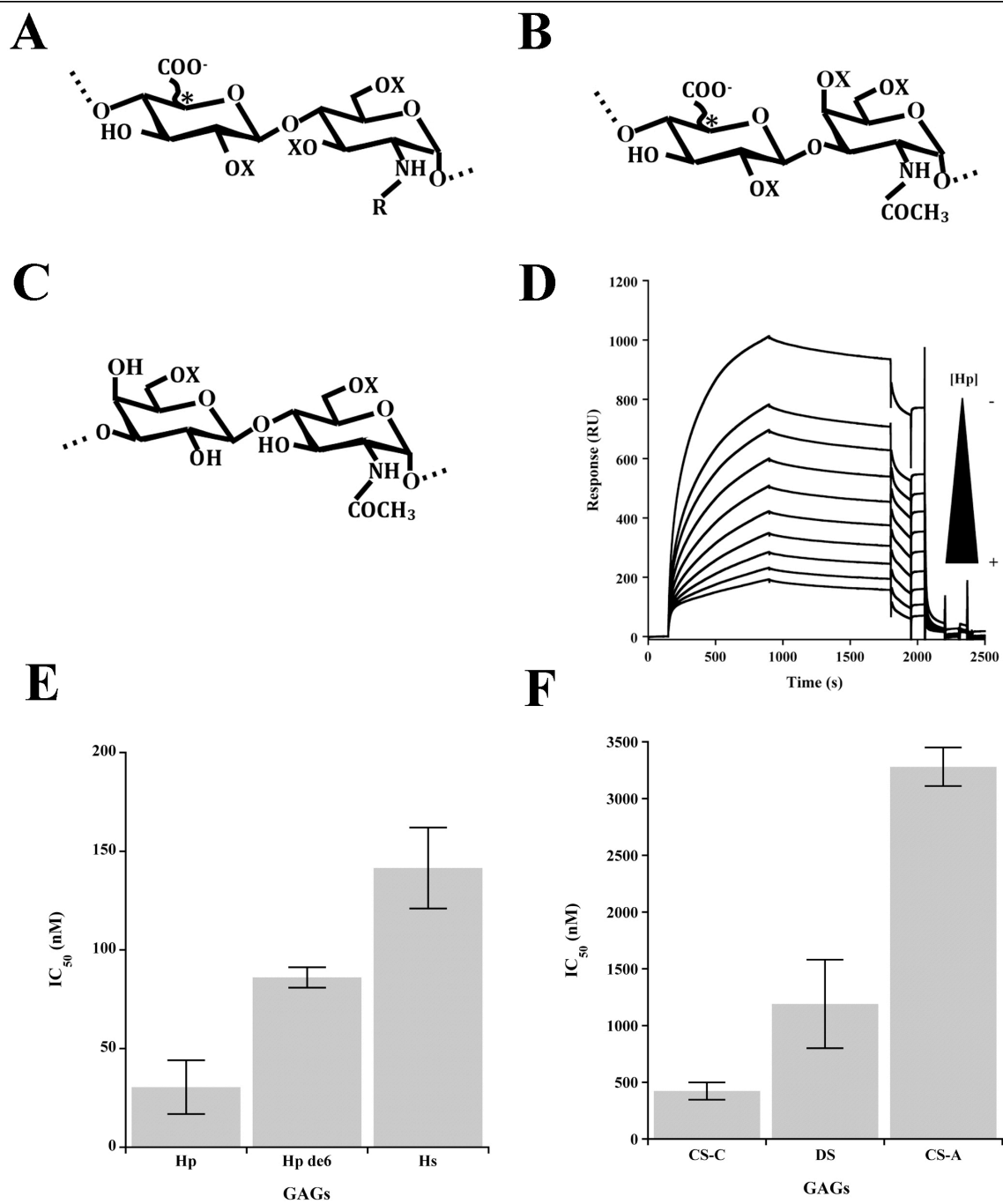
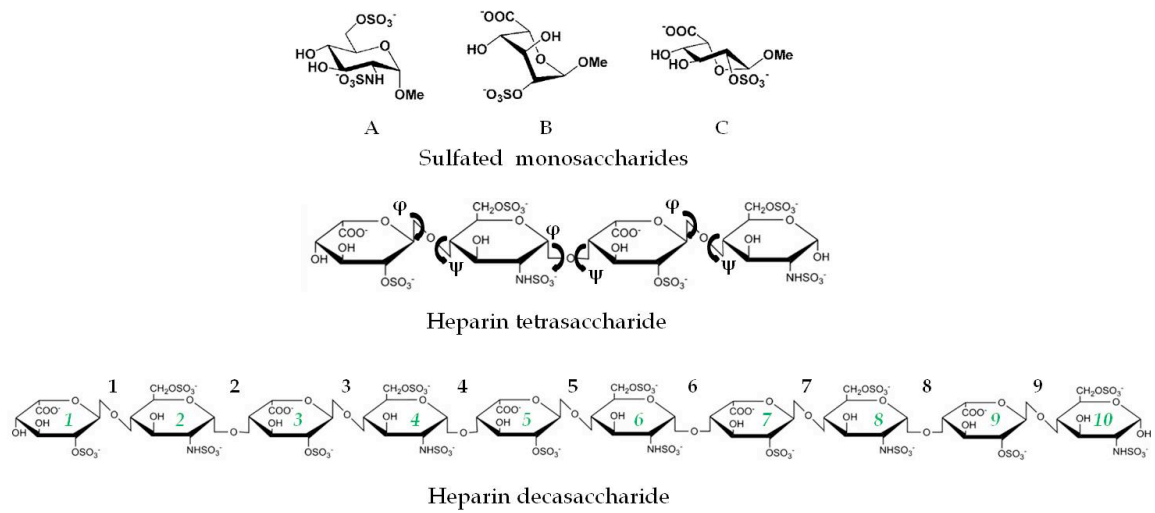


Figure 3





**Figure 4**



**Figure 5**

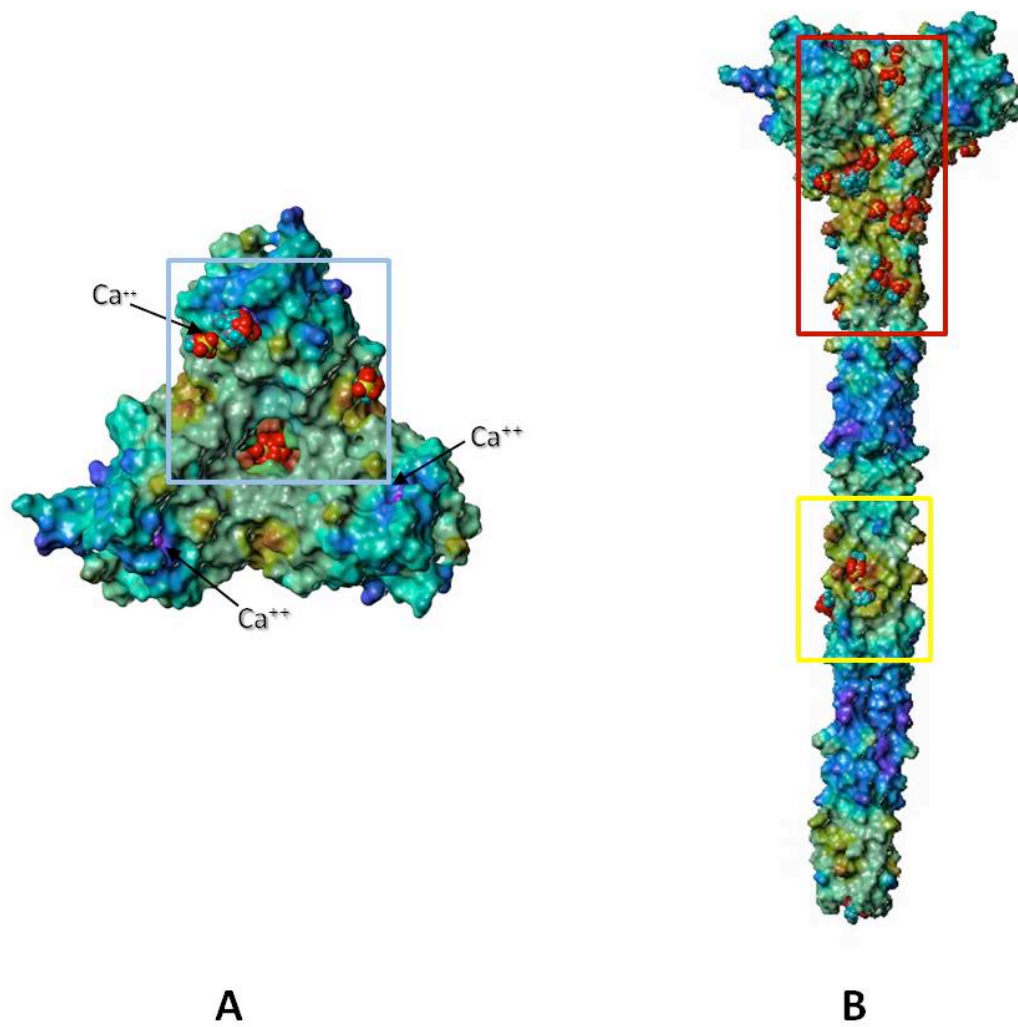
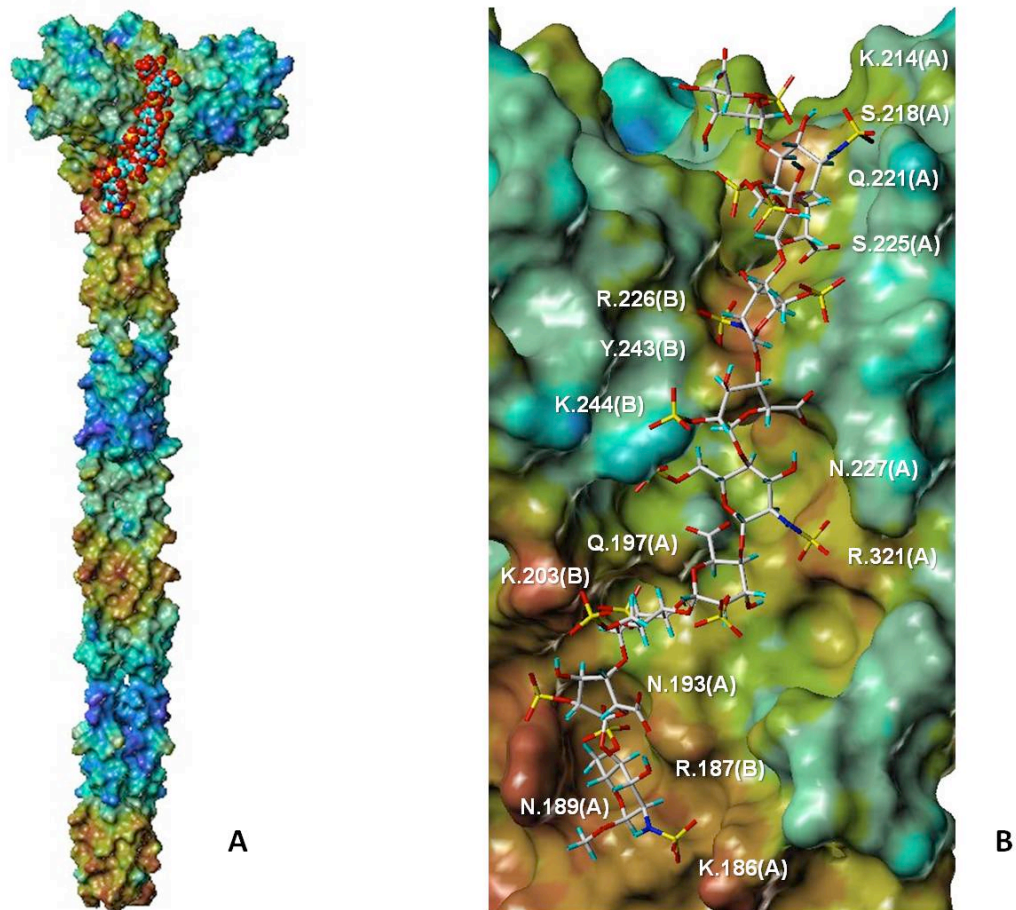


Figure 6



## 12.3 Résultats complémentaires

Au début de ce travail et après avoir observé la capacité de la Langérine à reconnaître l'héparine dans un site d'interaction qui ne dépend pas du calcium, plusieurs questions sont apparues :

- Quelle est la taille minimum requise de l'héparine pour que l'interaction soit possible (ou observable) ?
- Certains motifs saccharidiques sont-ils plus importants que d'autres dans le mode d'interaction, par exemple un sulfate indispensable à une position ?

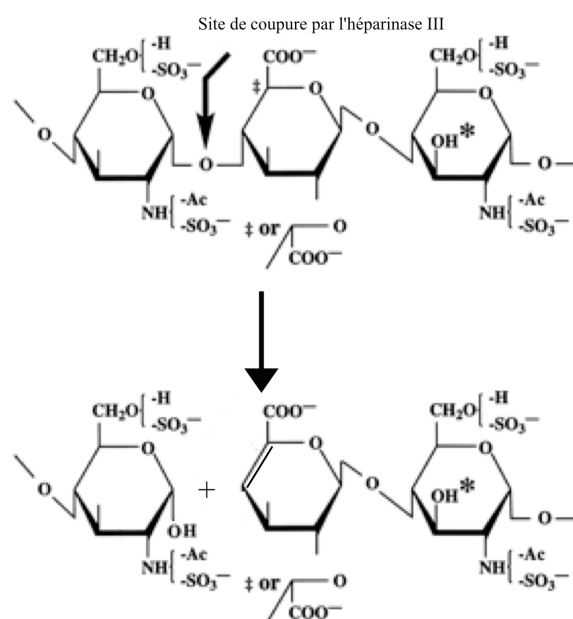


Figure 12.4 – Schéma de la digestion enzymatique de l'héparine par l'héparinase III. Modification de l'acide glucuronique par la présence d'une double liaison entre les carbones 4 et 5. [117]

### 12.3.1 Détermination de la taille minimum requise de l'héparine par SPR

Au début de l'étude de l'interaction de la Langérine avec les glycosaminoglycanes nous avons testé différentes tailles de glycosaminoglycanes obtenus par digestion enzymatique par l'héparinase III (l'expérience de digestion enzymatique et de purification des différents fragments ont été réalisées par Romain Vivès de l'équipe SAGAG de l'IBS) [108]. Cette étude a été réalisée par des expériences d'inhibition sur une surface préalablement fonctionnalisée avec successivement de la streptavidine puis de l'héparine (6 kDa) biotinylée et en utilisant comme piste de soustraction une surface fonctionnalisée seulement avec la

streptavidine. La Langérine, à une concentration fixe (100 nM), est incubée, ou non, avec des fragments d'héparine de différentes tailles et injectée sur les surfaces. Les fragments testés correspondent pour le plus petit à un hexasaccharide (dp6), constitué de trois fois l'unité répétée de l'héparine, et pour le plus grand à un dp18. L'enzyme responsable de la fragmentation de l'héparine clive uniquement la liaison  $\beta$ 1-4 liant le N-acetyl-glucosamine à l'acide glucuronique ou iduronique (Figure 12.4).

Pour cela tous les fragments sont constitués d'un nombre différent de ce disaccharide. L'acide glucuronique ou iduronique clivé est modifié par la présence d'une double liaison entre le carbone 4 et 5 due à la coupure enzymatique. Enfin, la méthode de purification des fragments ne différencie pas le nombre de sulfatation présent sur les fragments et on obtient donc différentes espèces dans chaque échantillon.

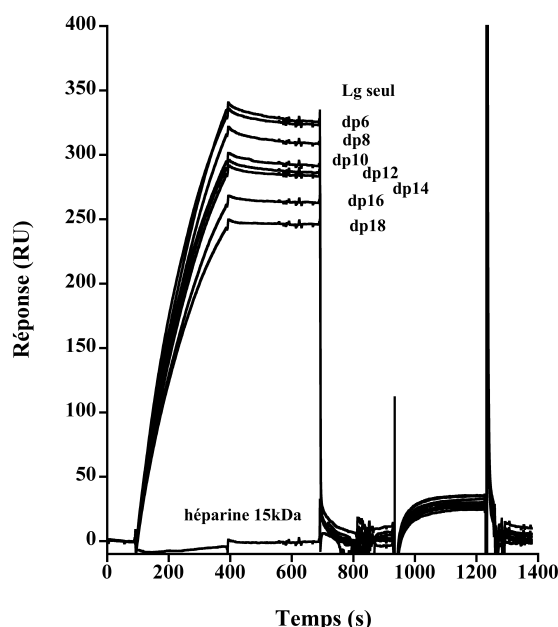


Figure 12.5 – Sensorgrammes de l'expérience d'inhibition de l'interaction de la Langérine avec l'héparine par différentes tailles de fragments d'héparine. Débit de 10  $\mu$ L/min. Injection de 50  $\mu$ L d'échantillons composés de Langérine à 100 nM et des différents fragments à 1  $\mu$ M. Les échantillons sont préparés dans le tampon de course (25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM  $\text{CaCl}_2$  et de 0,005% de surfactant P20).

Cette expérience a été réalisée deux fois consécutivement avec un ratio protéine/oligosaccharide de 1 :10 (Figure 12.5). Le volume injecté est limité (50  $\mu$ L) car la quantité des fragments d'héparine obtenue après purification est faible. Donc, dans un souci de faible consommation des fragments, nous avons dû diminuer la quantité de l'échantillon à injecter. Les contrôles positifs de l'interaction et de l'inhibition sont constitués respectivement de la Langérine seule et de la Langérine incubée avec de l'héparine 15 kDa. On observe de

faible inhibition pour les fragments d'héparine de toutes les tailles mais plus le fragment est long plus l'inhibition est bonne. Cette observation nous oriente à penser à un site d'interaction étendu de la Langérine pour les GAGs. L'objectif de cette expérience était d'observer un saut de l'inhibition par un des fragments qui nous aurait indiqué une taille minimum du fragment pour une bonne interaction. Or nous n'observons aucun saut de l'inhibition. Mais on peut observer que le dp6 n'inhibe presque pas, nous pouvons penser que cette taille d'héparine est la taille limite pour détecter une inhibition. Un pallier est observé pour les dp10-12-14 puis de nouveau une inhibition améliorée pour le dp16, cette observation laisserait à penser que le site peut être composé de deux régions qui nécessite 6 monosaccharides supplémentaires pour interagir dans les deux régions simultanément. Ceci n'est qu'une hypothèse car les expériences sont préliminaires. Pour observer plus précisément si une taille minimum est requise, il faudrait probablement injecter un plus grand volume pour se rapprocher le plus possible de l'équilibre qui permettrait d'observer, ou non, si le pallier d'inhibition existe vraiment. Mais comme évoqué précédemment, ces saccharides demandent beaucoup de temps pour les produire et l'obtention de l'équilibre par l'injection seule de la Langérine n'a jamais été obtenu même en injectant le volume maximum de 750  $\mu$ L.

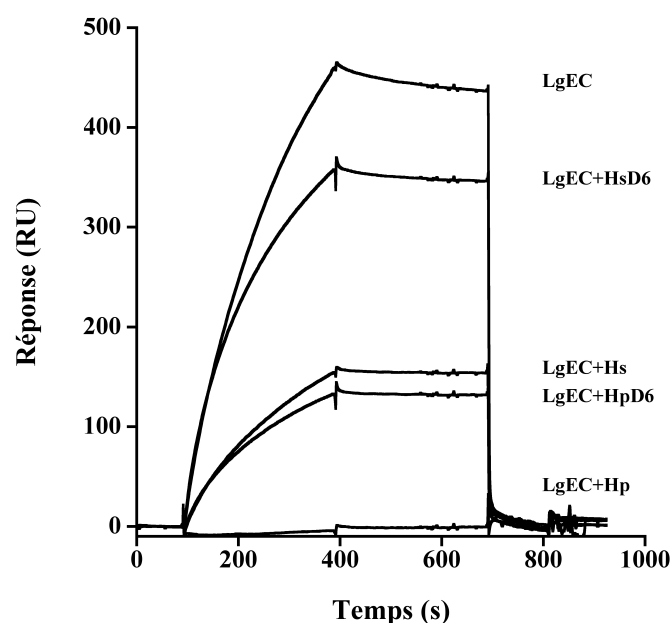


Figure 12.6 – Sensorgrammes de l'expérience d'inhibition avec les héparines et héparanes sulfates traités par HSulf2 ou non.

### 12.3.2 Le sulfate en C6 de la glucosamine important dans le mode d'interaction

Par la suite, J'ai eu l'opportunité de tester de l'héparine et de l'héparane sulfate 15 kDa traités par l'enzyme HSulf2. Cette enzyme permet d'enlever, le long du polysaccharide, les groupements sulfates en position 6 du disaccharide composé de l'acide iduronique sulfaté en 2 lié à un N-acetylglucosamine sulfaté en N2 et 6 (Idoa2S-GlcNS6S). Ce motif représente environ 5-7% des disaccharides. L'expérience a été réalisée sur la même sensor chip et dans les mêmes conditions expérimentales que l'expérience précédente (Figure 12.6).

La première observation montre la différence du pouvoir d'inhibition entre l'héparine et l'héparane sulfate suggérant que l'interaction est dépendante d'un motif très sulfaté. En effet, la différence majeure entre l'héparine et l'héparane sulfate est la proportion de groupement sulfate au sein du polysaccharide. Ensuite, on observe, dans les deux cas, que l'absence du groupement sulfate en position 6 du disaccharide Idoa2S-GlcNS6S diminue le pouvoir inhibant de ces molécules d'environ 30% (Figure 12.7). Cette observation suggère l'importance du motif disaccharidique fortement sulfaté dans le mode d'interaction de la Langérine pour l'héparine et l'héparane sulfate.

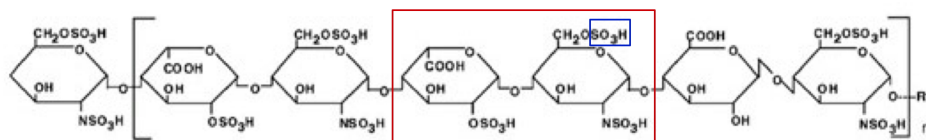


Figure 12.7 – Représentation de la chaîne d'héparine. Le cadre rouge met en évidence le disaccharide ciblé dans la désulfatation par l'enzyme HSulf2. Le cadre bleu montre le sulfate enlevé par HSulf2.

En conclusion de ces premières expériences complémentaires, nous avons constaté que pour observer une inhibition, dans des conditions limitées, un hexasaccharide est la taille limite. En effet, l'affinité de la Langérine pour l'hexasaccharide n'est pas assez forte, comparé à son l'affinité pour une héparine de 6 kDa, pour observer une inhibition. De plus, on a vu que plus l'héparine est longue plus l'inhibition est importante. Ces observations suggèrent un site étendu d'interaction entre la Langérine et l'héparine mais aussi la possibilité que des longues chaînes d'héparine peuvent interagir avec plusieurs molécules de Langérine ECD expliquant ainsi l'amélioration de l'inhibition par l'allongement de la chaîne saccharidique. Ensuite, le dissaccharide trisulfaté joue un rôle non discutable dans le mode d'interaction et en particulier le sulfate en position C6. En effet, le traitement par HSulf2 de l'héparine, ou l'héparane sulfate, modifie significativement l'affinité de la Langérine pour ces sucres comme en atteste la diminution de l'inhibition représentée dans

la figure 12.6. Ainsi, les motifs fortement sulfatés le long de la chaîne des GAGs sont à l'origine de ce nouveau mode d'interaction de la Langérine.

### 12.3.3 Expérience d'inhibition de l'interaction Langérine/gp120 ou BSA-Man par l'héparine

Dans ces expériences, nous avons voulu savoir si l'héparine était tout de même capable d'inhiber l'interaction de la Langérine avec gp120 ou la BSA-Man. Sur des surfaces fonctionnalisées avec gp120 ou la BSA-Man, nous avons injecté, avec une concentration fixe en protéine, différents ratio héparine/protéine. La première surprise a été d'observer non pas une inhibition mais une augmentation du signal correlant à l'augmentation du ratio héparine/protéine (Figure 12.8). Cette observation indique que ces sites d'interactions sont différents, cela a déjà été observé par les expériences d'interaction de la Langérine avec l'héparine en tampon calcium et EDTA, et qu'ils ne sont probablement pas proches. L'héparine n'interagissant pas avec les gp120 de tropisme R5, l'interaction du complexe avec gp120 n'est donc pas médié par l'héparine.

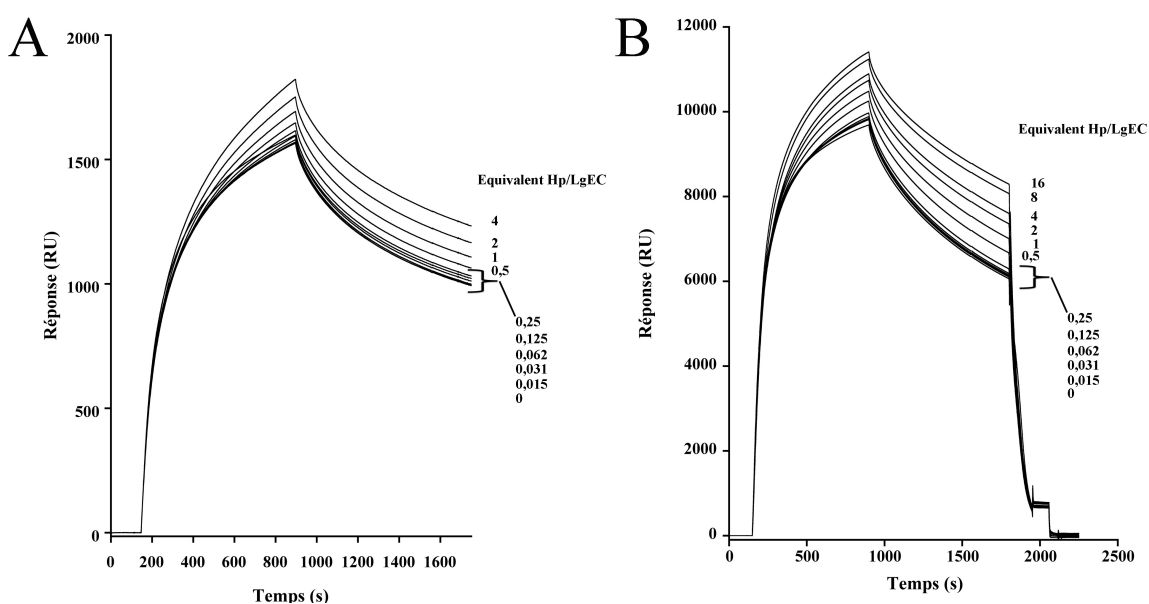


Figure 12.8 – Sensorgrammes de l'expérience d'inhibition de l'interaction Langérine/gp120 (A) ou BSA-Man (B) par l'héparine.

Nous observons une amélioration de la réponse par l'augmentation du ratio héparine/Langérine aussi bien pour l'interaction avec gp120 et la BSA-Man. Nous pouvons donc conclure que les sites d'interaction de la Langérine pour les « High Mannose » et pour les glycosaminoglycanes sont indépendants et cumulables.



### 12.3.4 Résultats préliminaires sur la formation de complexe macromoléculaire formé de Langérine et d'héparine : vers une description des granules de Birbeck.

On a vu précédemment que la Langérine est indispensable pour la formation des granules de Birbeck. Plusieurs études l'ont démontré. En effet, les investigations menées par Jenny Valladeau montrent que l'expression de la Langérine dans des fibroblastes de souris est suffisante à la formation de granule de Birbeck[104]. Par la suite, des mutations ponctuelles dans le domaine CRD de la Langérine induisent la perte ou une modification drastique des granules de Birbeck montrant l'importance de l'intégrité de la Langérine pour la formation des BGs [105, 107]. Ensuite, l'étude menée par l'équipe en collaboration avec Jenny Valladeau a confirmé l'importance de la Langérine dans cet organelle particulier, d'une part, et affiné le rôle joué de certains domaines de la Langérine pour la formation des BGs [100]. Dans cette étude, un modèle d'organisation des BGs ainsi que sa formation a été proposé (Figure 5.3).

Ce modèle suggère que la Langérine doit être capable de s'auto-assembler et/ou est qu'il y a une matrice centrale saccharidique qui permet de maintenir cette organisation ?

Dans l'hypothèse d'une matrice centrale, les glycosaminoglycans sont des bons candidats car ce sont des polysaccharides très longs, linéaires et abondants dans la matrice extracellulaire. De plus l'interaction de la Langérine avec les GAGs n'empêche pas la Langérine d'interagir par son site canonique avec des sucres de type « High Mannose ».

Ans, pour tester ces hypothèses, des expériences préliminaires ont été réalisées par microscopie électronique et par SPR.

#### 12.3.4.1 Microscopie électronique

Des tests préliminaires de microscopie électronique à transmission ont été réalisés sur la Langérine ECD seule ou en complexe avec de l'héparine 15 kDa. Les échantillons ont été préparés par la méthode de coloration négative, donc les protéines apparaissent plus claires que les zones recouvertes par le contrastant (Figure 12.9).

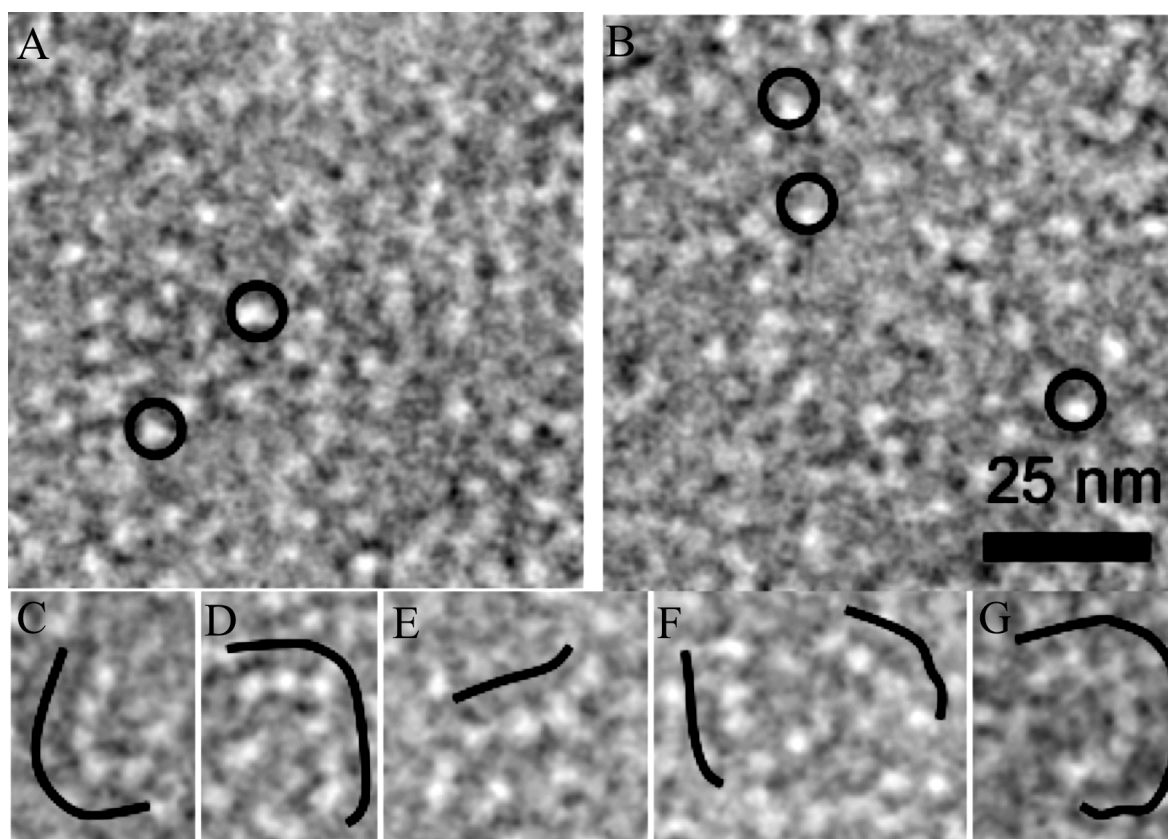


Figure 12.9 – Micrographies de la Langérine ECD seule (A,B) ou en complexe avec de l'héparine (C, D, E, F, G). A-B : LgECD seule. C-D-E-F-G : LgECD en complexe avec l'héparine 15 kDa pour une ratio 1 :1. L'échelle est identique pour toutes les micrographies.

Les premières micrographies de la protéine seule nous permettent d'observer les domaines CRDs mais pas le domaine neck (A, B). Cela est due à l'organisation du coiled-coil qui par conséquent impose une structure étendue de faible diamètre que l'on ne peut observer par cette technique. Lorsque l'on ajoute de l'héparine à l'échantillon, on observe des densités continues et linéaires ou incurvées (Figure 12.9, C-D-E-F-G). Ces densités linéaires sont composées d'une succession de petites densités de même taille que celles observées pour la protéine seule. Attention on peut observer aussi des densités linéaires pour la protéine seule, mais celles-ci possèdent des petites densités nettement plus espacées que dans les échantillons possédants de l'héparine. Ces résultats préliminaires nous permet de suggérer que l'héparine est capable de former des chapelets de Langérine le long de sa chaîne.

Pour pouvoir aller plus loin, il faudrait être capable de visualiser le côté N-terminal de la protéine, c'est à dire l'extrémité du coiled-coil. Pour cela, nous avons pensé utiliser une autre construction de la Langérine ECD qui possède une étiquette « Strep Tag II » à l'extrémité N-terminale. Cette étiquette nous permettrait, peut être, d'observer l'extrémité

N-terminale lorsque celle-ci est en interaction avec la streptactine. La streptactine reconnaît spécifiquement le « Strep Tag II » et est une protéine globulaire tétramérique d'environ 60 kDa dérivant de la streptavidine. La streptactine fixée à l'extrémité N-terminale de la Langérine serait donc détectable. Cela pourrait nous donner des informations sur la disposition des protéines le long de la chaîne d'héparine. En effet on pourrait discriminer si l'extrémité N-terminale est perpendiculaire à la chaîne d'héparine ou le long de celle-ci afin de présumer d'un rôle possible de ces GAGs dans la structure du granule de Birbeck mais ceci est purement spéculatif pour le moment.

#### 12.3.4.2 Propriétés d'assemblage de la Langérine

Afin d'étudier la capacité d'interaction de la Lg-ECD sur elle-même par SPR, des surfaces orientées fonctionnalisées avec de la Langérine ECD ont été réalisées. Pour cela nous avons utilisé une construction de la Langérine ECD présentant une étiquette « Strep Tag II » à l'extrémité N-terminale, la Langérine S-ECD (Figure construction). J'ai donc premièrement immobilisé de la streptactine sur deux canaux d'une sensor chip CM4, puis capturé la Lg-S-ECD sur la seconde. Mais cette deuxième immobilisation ne perdure pas dans le temps sans une activation préalable par l'EDC/NHS afin de lier covalamment les deux protéines. Ceci est probablement dû à une affinité de la streptactine pour le « Strep Tag II » pas suffisante pour ce type d'expérience (affinité de l'ordre du  $\mu\text{M}$ ). Cette expérience de fin de thèse a été réalisée une seule fois (Figure 12.10).

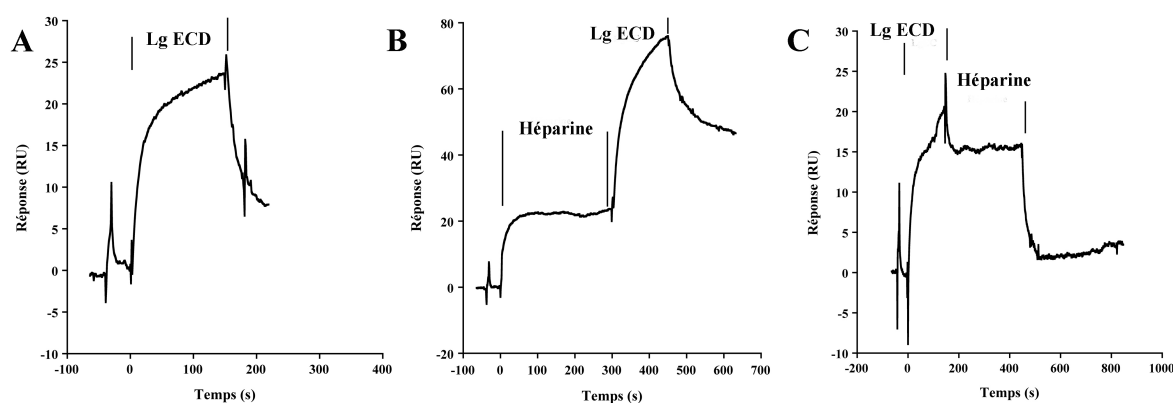


Figure 12.10 – Sensorgramme de l'expérience d'injection successive de la Langérine et de l'héparine. A : Injection de la Langérine seule (50  $\mu\text{L}$  à 500 nM). B : Injection de l'héparine (100  $\mu\text{L}$  à 100 nM) puis de la Langérine (50  $\mu\text{L}$  à 500 nM). C : Injection de la Langérine puis de l'héparine. Le débit est de 20  $\mu\text{L}/\text{min}$  dans le tampon calcium (25 mM Tris pH 8, 150 mM NaCl, 4 mM  $\text{CaCl}_2$  et 0,005% P20).

Premièrement, nous observons une interaction de la Langérine avec la Langérine (Figure 12.10, A) de l'ordre de 25 RU. Cette interaction est faible et surtout elle possède

une dissociation rapide. Mais lorsque que l'on injecte tout d'abord de l'héparine avant la Langérine, on observe bien sûr l'interaction de l'héparine avec la protéine immobilisée (Figure 12.10, B), qui donne une réponse d'environ 20 RU, mais surtout on observe une réponse 2 fois plus importante de l'interaction Langérine comparé à celle sans injection préalable d'héparine. Nous observons peut être ici un effet de pontage de l'héparine vu en microscopie électronique par SPR. Enfin, lorsque l'on injecte l'héparine après la Langérine, on observe une stabilisation du signal qui est dû probablement à une compensation de la dissociation de la Langérine par l'association de l'héparine (Figure 12.10, C).

L'effet de pontage de Langérine par l'héparine observé en microscopie électronique est peut être visualisé dans l'expérience de SPR (Figure 12.10, B). Ces investigations sont que préliminaires, il est nécessaire pour affirmer quoique se soit de refaire les expériences en augmentant la quantité de Langérine S-ECD greffée pour obtenir des signaux plus importants.

Enfin, en conclusion de ce chapitre, il me semble indispensable d'effectuer des expériences de biologie cellulaire pour affirmer ou infirmer le rôle des GAGs dans la formation des BGs par le pontage de Langérine. Des premières expériences menées en collaboration ne nous ont pas permis de répondre à cette question dû à une incertitude sur l'efficacité d'un des produits utilisés pour empêcher la production de GAGs par la cellule.

# Chapitre 13

## Caractérisation structurale de l'interaction de la Langérine avec des sucres sulfatés

### 13.1 Introduction

Pour aborder le nouveau mode d'interaction de la Langérine pour les glycosaminoglycanes, une étude structurale est nécessaire pour confirmer ou infirmer les suggestions apportées par la modélisation. Les approches envisagées pour caractériser l'interaction de la Langérine avec des saccharides sulfatés ont été la RMN STD et la cristallographie. La RMN a été abordée au travers d'une collaboration (Jesus Angulo et Pedro Nieto, IIQ, CSIC Séville) et l'approche par cristallographie a été réalisée par mes soins à l'IBS.

Nous avons montré dans le chapitre précédent que l'interaction de la Langérine avec l'héparine est dépendante de l'oligomérisation de la protéine. Au démarrage de cette partie structurale de mon travail, la structure cristallographique d'une forme oligomérique de la Langérine, publiée en 2010 par Feinberg et al [38], n'existait pas. Donc ma première approche a été d'essayer de cristalliser la Langérine ECD. Mais, on a vu aussi que le domaine neck semble très flexible, alors nous avons décidé d'utiliser des composés « pontants », développés par l'équipe de Richard Kahn, afin de cristalliser la forme trimérique de la Langérine et ainsi disposer de l'interface entre protomères nécessaire au site d'interaction avec l'héparine. Nous avons donc tenté de co-cristalliser le domaine ECD avec le dipicolinate d'euporium (Fig 13.1 A). Cette molécule possède des propriétés intéressantes pour la cristallisation de macromolécules car elle a la faculté de pouvoir « ponter » des molécules grâce à la fonction dipicolinate ainsi que d'obtenir des phases expérimentales grâce à la présence de l'euporium (Fig 13.7 B). Afin d'obtenir des pistes de cristallisation, les échantillons ont été envoyés au robot de cristallisation en nano-goutte du PSB

(« Partnership for Structural Biology »).

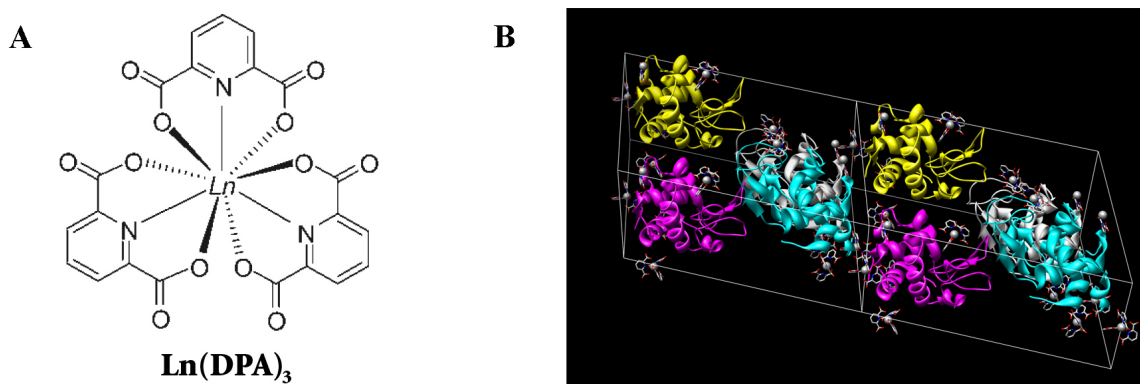


Figure 13.1 – Structure du tris-dipicolinate de lanthanide (A) et de la capacité de pontage de cette molécule observée dans le cas du lysosyme [81].

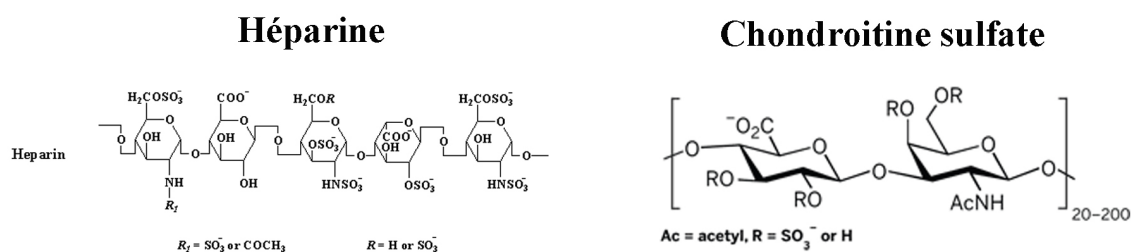
La cristallisation avec une longue chaîne d'héparine s'annonçait difficile. Ainsi les essais de cristallisation du domaine ECD ou CRD de la Langérine incubé avec de l'héparine 15 kDa n'ont pas donné de résultats. La difficulté majeure avec ce type d'oligosaccharide est l'hétérogénéité de la chaîne oligosaccharidique. Nous avons donc décidé d'utiliser des fragments commerciaux de saccharides provenant soit de l'héparine soit des chondroitines sulfates (Figure 13.2).

Notre choix s'est porté sur des disaccharides sulfatés. Nous avons aussi fait le choix d'utiliser des disaccharides et non des oligosaccharides plus long dû à l'hétérogénéité des échantillons obtenus par digestion enzymatique. En effet, les saccharides étudiés sont obtenus par digestion enzymatique d'une longue chaîne de glycosaminoglycanes, donc pour chaque motif oligosaccharidique obtenu, il existe une large diversité de sucre selon le degré et l'emplacement des groupements sulfates et plus l'oligosaccharide est long plus le nombre d'espèces augmente. Cette digestion provoque aussi au niveau de l'extrémité non réductrice des disaccharides une double liaison entre les carbones 4 et 5 de l'acide iduronique empêchant une interaction directe dans le site calcium du fait de l'absence du groupement hydroxyle en position C4. L'interaction avec l'héparine étant la plus importante, nous nous sommes plus particulièrement focalisés sur les saccharides de la série héparine avec 4 disaccharides et 3 monosaccharides. La différence entre tous ces saccharides est la présence ou non de groupements sulfates à différentes positions. Pour la série chondroïtine sulfate, deux disaccharides ont été utilisés, un disaccharide non sulfaté et celui sulfaté en position C6 sur le GalNac. Nous avons utilisé des disaccharides car ils représentent le motif de base répété le long de la chaîne osidique. De plus, la liaison osidique utilise soit l'hydroxyle en C4 soit l'hydroxyle en C3 qui sont nécessaires pour la liaison dépendante de l'atome de calcium. On peut donc espérer favoriser une fixation dans le site héparine

# CHAPITRE 13. CARACTÉRISATION STRUCTURALE DE L'INTERACTION DE LA LANGÉRINE AVEC DES SUCRES SULFATÉS

par rapport au site calcium si la cristallisation est possible.

**A**



**B**

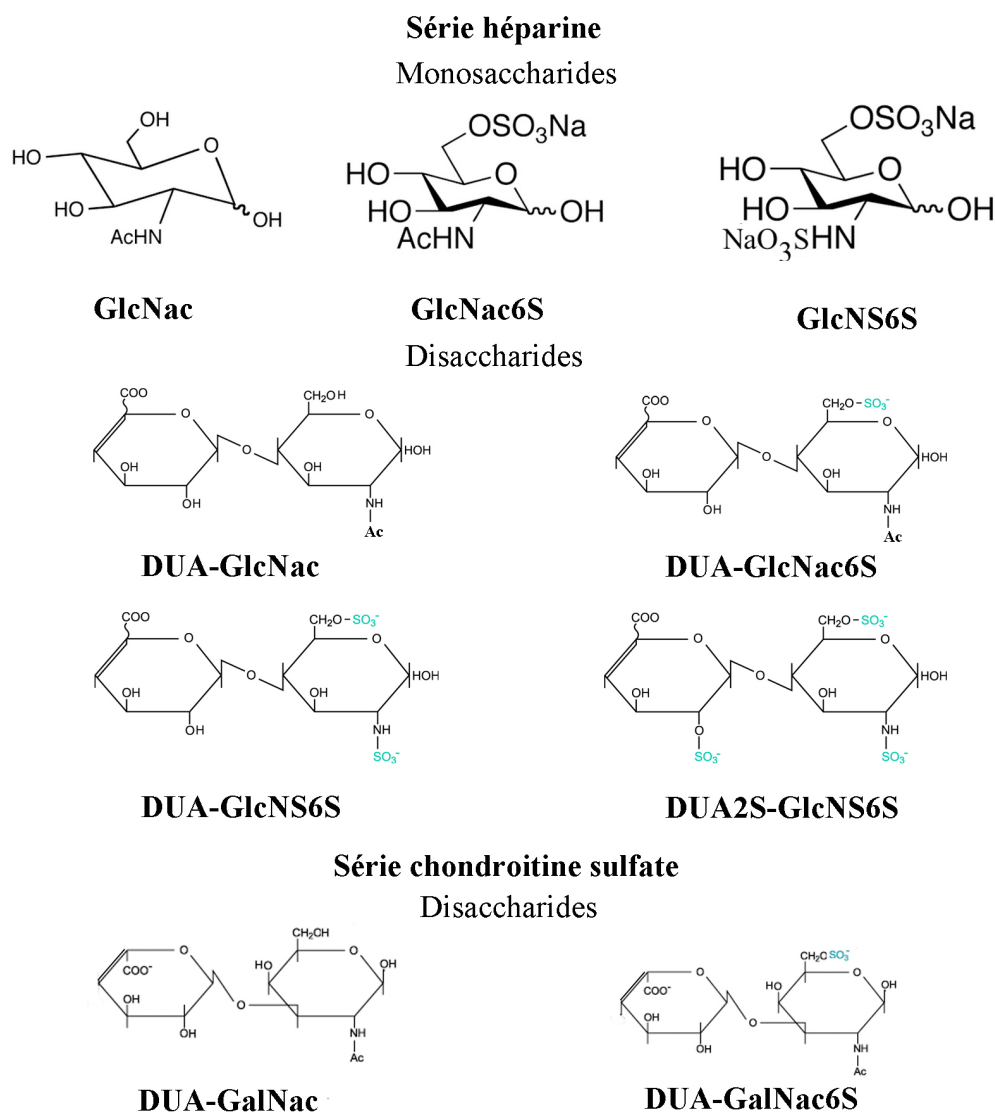


Figure 13.2 – Structures des saccharides utilisés.

Des conditions sans et avec des fragments de glycosaminoglycanes ont été testées en co-cristallisation avec la Langérine ECD. Certains sucres utilisés ont aussi été testés en co-cristallisation avec la Langérine CRD. Pour chaque échantillon, 540 conditions commerciales de cristallisation ont été testées avec 6 visualisations par condition (24 h, 72 h, 1 semaines, 2 semaines, 5 semaines et 10 semaines). Au total environ 39 000 images ont été visualisées pour au final ne trouver qu’une seule condition positive. Cette condition de cristallisation a permis d’observer en 72 h de tous petits cristaux. Elle a été obtenue à partir d’un des tests avec la Langérine CRD, avec le disaccharide sulfaté de la série des chondroitines sulfates (DUA-GalNac6S). Aucune co-cristallisation n’a été obtenue pour les autres disaccharides. A partir de cette piste, j’ai donc réalisé une première série de conditions proches de celle trouvée en nano-goutte. Mais le passage des nano-gouttes en micro-gouttes est souvent délicat car la diffusion de vapeur est différente. Pour cette raison, une longue procédure de cristallogenèse a été réalisée jusqu’à obtenir une seule condition possédant un unique cristal au bout de 72 h. Pour obtenir ce cristal, il a fallu augmenter drastiquement la concentration en sulfate de magnésium (de 1 M en nano-goutte à 1,4 M en micro-goutte) et changer légèrement le pH du tampon (de 6,5 à 7). Ce cristal, obtenu en fin de thèse, a été analysé par diffraction des rayons X sur la ligne FIP (BM30a) de l’ESRF mais aucune diffraction n’a été observée. Ceci est probablement due à une quantité de solvant élevée dans le cristal qui était très fragile indiquant qu’il était constitué de protéine et non d’un sel. De plus, un cristal de sel aurait diffracté contrairement à celui-ci. Au final, aucun disaccharide n’a permis d’obtenir des informations structurales sur le mode de fixation des glycosaminoglycanes avec la Langérine.

Outre les expériences de co-cristallisation avec les sucres, les essais de cristallisation de la Langérine ECD seule ou en présence de dipicolinate d’euprium ont conduit à l’obtention de cristaux. Ceux-ci apparurent dans des conditions similaires à celle décrite pour la Langérine CRD. Il s’est avéré, après analyse, que ces cristaux étaient constitués du domaine CRD provenant d’une protéolyse spontanée dans la goutte (voir section 13.3).

N’obtenant aucun résultat pour la cristallisation dans le site héparine, nous avons décidé de réorienter l’étude structurale. A ce moment, la caractérisation de l’interaction de la Langérine avec des sucres sulfatés dans le site calcique n’était pas encore publiée. Nous avons donc voulu caractériser ce type d’interaction pour des monosaccharides sulfatés (Figure 13.2). Nous avons utilisé les cristaux précédemment obtenus, notamment ceux créés à partir de la Langérine ECD en présence d’euprium, pour déterminer les structures en présence des monosaccharides par la méthode de trempage. Cette méthode consiste à faire diffuser les saccharides dans les canaux de solvant présents dans les cristaux. Les monosaccharides sulfatés peuvent interagir dans le site calcium par la présence des groupements hydroxyles en C3 et C4 en position équatoriale dans le cas des monosaccharides dérivés du



GlcNac. Pour cette approche, les conditions de cristallisation de la protéine sont connues, il est donc aisé de produire un grand nombre de cristaux. La seule difficulté est le temps de cristallisation puisque deux mois sont nécessaires pour obtenir des cristaux d'une taille suffisante (50  $\mu\text{m}$  dans la plus petite dimension). Les trempages avec tous les monosaccharides ont été réalisés avec des cristaux de Langérine CRD préalablement obtenues à partir de langérine ECD (l'obtention et l'analyse des cristaux à partir de Langérine ECD sont expliquées dans la section cristallographie). Les données de diffraction ont été enregistrées sur la ligne de lumière FIP (BM30a) de l'ESRF. Au cours de l'enregistrement des données, une analyse rapide des données, sans avoir recours à un remplacement moléculaire, nous a permis de déterminer la présence de saccharide ou non. Ceci nous a permis d'observer des densités électroniques au niveau du site canonique pour tous les monosaccharides. Cette première analyse rapide nous a donc permis d'identifier les jeux de données à analyser en détails.

Parallèlement à la détermination des structures, nous nous sommes attelé à caractériser l'affinité de ces différents ligands pour la Langérine par SPR.

## 13.2 Etude fonctionnelle par Résonance plasmonique de surface

Les monosaccharides étant de masse faible, l'étude de SPR n'est pas aisée. En effet, par SPR l'intensité du signal mesuré est proportionnelle à la masse de la molécule qui se lie à la protéine dans le cas où la protéine est greffée sur la surface. N'ayant pas la possibilité de greffé les sucres, il est donc difficile d'estimer un  $K_d$ , mais nous pouvons greffé la BSA-Man (BSA couplé avec des trimannoside, environ 12 sites de glycosilations sur la protéine) qui interagit dans le site calcium de la Langérine. Donc par des expériences d'inhibition, nous pouvons évaluer la capacité d'inhibition de chaque sucre dans l'interaction de la Langérine avec le mannose exposé par la BSA. Les expériences d'inhibition ont été réalisées à la fois sur une surface fonctionnalisée avec la BSA-Man (Figure 13.3), mais aussi sur une surface immobilisée d'héparine pour confirmer la non-inhibition dans le site héparine (surface fonctionnalisée tel que décrit dans le chapitre 12). Parallèlement à l'étude sur la Langérine, nous avons aussi évalué l'affinité de ces ligands pour DC-SIGN, une autre lectine de type-C étudiée au laboratoire, qui est structuralement et fonctionnellement assez proche de la Langérine notamment dans sa capacité à reconnaître le VIH. En effet, en marge de ce travail centré sur la Langérine, l'équipe développe des inhibiteurs spécifiques de DC-SIGN. Ainsi, toute identification de modification sur les ligands conduisant à une spécificité d'une lectine par rapport à l'autre présente un intérêt. Nous avons donc réalisé des expériences d'inhibition par compétition avec la Langérine et DC-SIGN. Sur une surface de type CM4

(GE Healthcare sensor chip), nous avons immobilisé respectivement une quantité similaire de BSA et BSA-Man, environ 4600 RU, sur la piste référence et sur la piste 2 (Figure 13.3).

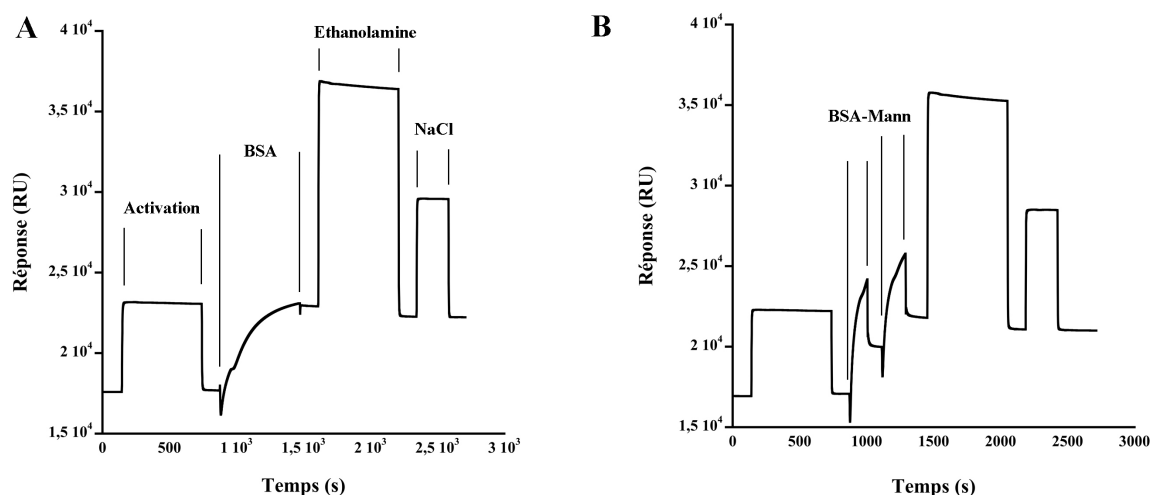


Figure 13.3 – Sensorgramme des immobilisations de BSA (A) et BSA-Man (B).

Les premiers résultats nous ont montré que ces saccharides ne sont pas capables d'entrer en compétition avec l'héparine pour l'interaction avec la Langérine (données non montrées). Par contre, ils peuvent, par compétition, dissocier un complexe Langérine/BSA-Man formé sur une surface biacore. Cette observation suggère que la présence des hydroxyles libres en C3 et en C4 permet une reconnaissance dans le site calcique malgré la présence du groupement sulfate. En effet, ces hydroxyles libres sont présents dans tous les sucres étudiés dans cette partie. Les groupements sulfates ne sont donc pas la condition unique pour une interaction dans le site héparine indépendant du calcium. Ce résultat montre que l'interaction dans le site héparine est dépendant, non seulement des groupements sulfates, mais surtout du fait de la non disponibilité de l'hydroxyle en C4 libre qui est utilisé dans la liaison osidique dans l'héparine.

Les questions résultantes de cette observation sont les suivantes :

- Est-ce que ces groupements sulfates portés par les GlcNac apportent une amélioration de l'interaction comparée à un monosaccharide non sulfaté ?
- Le groupement sulfate induit-il une spécificité des sucres pour la Langérine vis-à-vis de DC-SIGN ?

Pour répondre à ces questions, nous avons déterminé les  $IC_{50}$  de différents saccharides afin d'établir l'importance des groupements sulfates. En effet, dans cette étude par SPR, nous avons utilisé les monosaccharides sulfatés ou non dérivés du GlcNac (GlcNac, GlcNac6S et GlcNS6S) ainsi qu'un trisaccharide Trila qui est décrit dans la section 13.4. Nous avons injecté à une concentration fixe, 100 nM, de Langérine ou DC-SIGN seul ou en

présence des différents saccharides à des concentrations variables allant de 0 à 24 mM pour le GlcNac et jusqu'à 6 mM pour le GlcNac6S et GlcNS6S (Tableau 8.1). Nous observons une diminution de la réponse maximum sur les sensorgrammes indiquant le pouvoir d'inhibiteur compétitif de ces molécules pour les interactions dépendantes du site calcium. A partir de ces sensorgrammes, nous sommes capables d'estimer la concentration de sucre qui permet d'inhiber 50% de la réponse maximale sans compétiteur ( $IC_{50}$ ) (Figure 13.4).

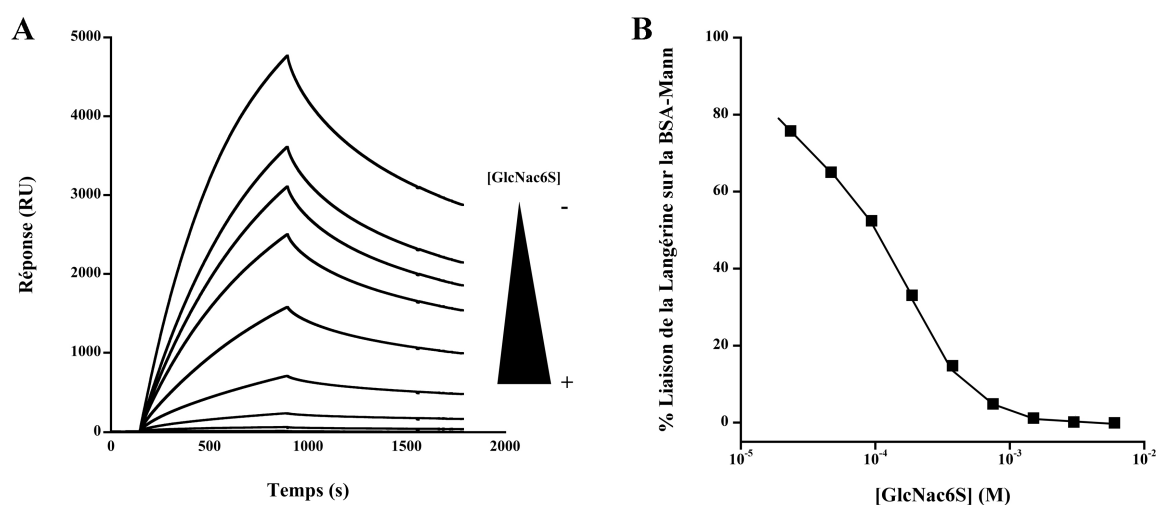


Figure 13.4 – Sensorgrammes d'une expérience d'inhibition avec le GlcNac6S (A) et la représentation graphique permettant d'estimer l' $IC_{50}$  (B).

A partir des différentes  $IC_{50}$  estimées pour les sucres étudiés, on constate un  $IC_{50}$  de l'ordre du mM dans le cas du GlcNac aussi bien pour la Langérine que pour DC-SIGN, cette observation est en adéquation avec la faible affinité attendue pour un monosaccharide non sulfaté (Figure 13.5).

Par contre, on observe une amélioration d'un facteur 10 de la reconnaissance de la version sulfaté de ces monosaccharides pour la Langérine et non pour DC-SIGN (Figure 13.5). En effet, on observe une augmentation significative de la valeur d' $IC_{50}$  du GlcNac6S par rapport au GlcNac pour DC-SIGN indiquant que l'interaction de DC-SIGN avec ce type de sucre est moins favorable que le sucre non sulfaté. Pour la Langérine, l' $IC_{50}$  est diminué d'un facteur 10 pour les sucres sulphatés de type GlcNac6S et GlcNS6S montrant la forte spécificité de la Langérine pour ce type de sucre. Effectivement, la différence des  $IC_{50}$ , du GlcNac6S pour la Langérine et DC-SIGN, est d'un facteur 100. La charge négative en position C6 entraîne donc un différentiel de spécificité en faveur de la Langérine. Ces résultats montrent clairement que le groupement sulfate important pour l'augmentation de l'affinité vis-à-vis de la Langérine est celui porté par le carbone en position C6 du glucosamine. Puisque l' $IC_{50}$  du GlcNS6S, doublement sulfaté, n'est pas significativement

amélioré par rapport au GlcNac6S.

Pour le trisaccharide Tri1a, l' $IC_{50}$  est globalement similaire à celui du monosaccharide non sulfaté où seul les hydroxyles en C3 et C4 interviennent dans l'interaction. Ceci est surprenant car le trisaccharide interagit probablement par son extrémité non-réductrice, qui est la seule à présenter des hydroxyles en C3 et C4 libres pour une interaction avec le calcium. Cette extrémité est un glucosamine sulfaté en C6, or nous avons montré que pour le monosaccharide correspondant ce sulfate améliore d'un facteur 10 l' $IC_{50}$ . L'absence de cet effet pour le Tri1a suggère l'existence de contraintes stériques spécifiques dans le contexte du trisaccharide.

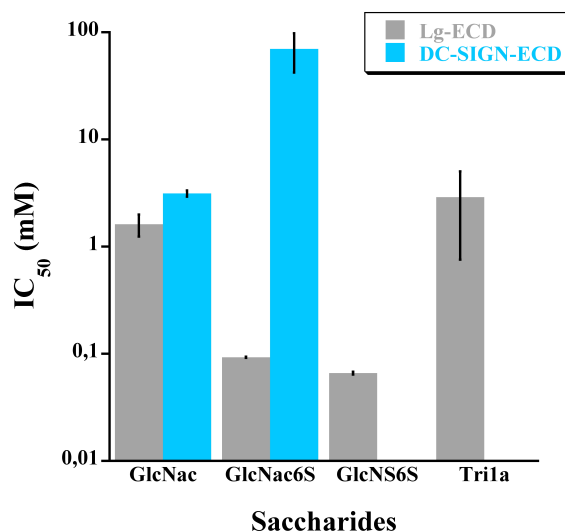


Figure 13.5 – Histogramme des  $IC_{50}$  des sucres de type GlcNac, GlcNac6S, GlcNS6S et Tri1a pour inhiber l'interaction de la Langérine ou DC-SIGN avec la BSA-Man.

Cette étude par SPR permet de répondre aux questions précédentes :

- L'interaction avec des saccharides sulfatés de type GlcNac est spécifique de la Langérine et non de DC-SIGN. En se rapportant aux structures des domaines CRD de ces deux lectines, on observe la présence de deux lysines à proximité du site calcium de la Langérine qui ne sont pas conservées chez DC-SIGN (Figure 13.6). La présence de ces deux lysines peut expliquer la spécificité de reconnaissance de la Langérine comparé à DC-SIGN pour les sucres sulfatés.
- Le groupement sulfate en position 6 apporte une amélioration significative de l'interaction avec la Langérine.

Nous avons étudié les structures de la Langérine en complexe avec les sucres afin de compléter cette analyse.

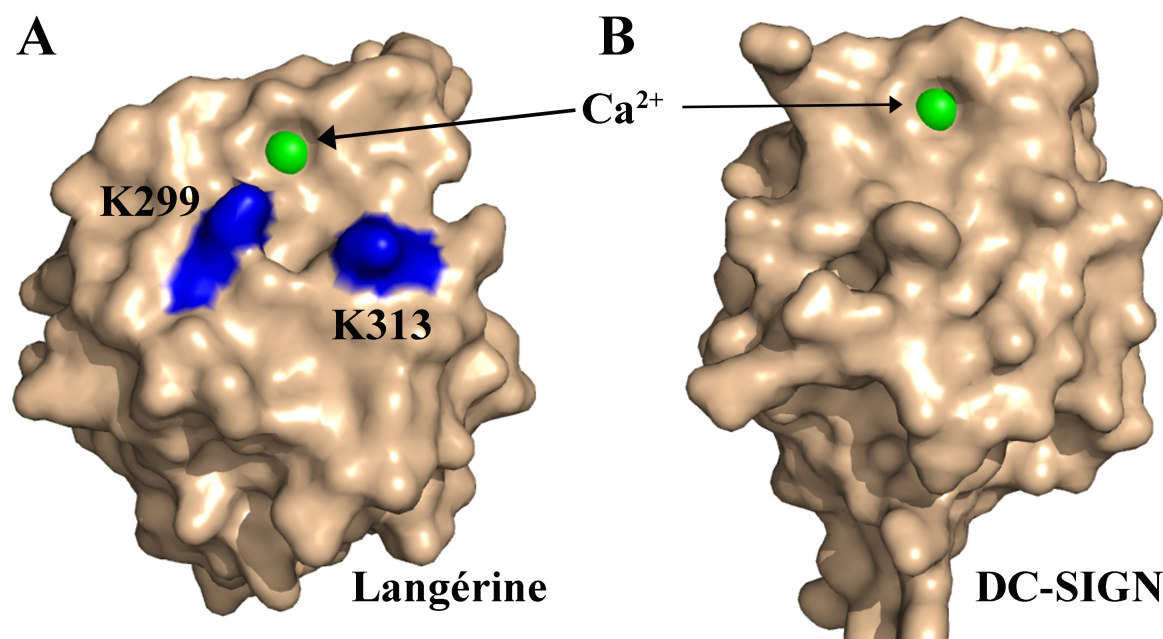


Figure 13.6 – Structures cristallographiques de la Langérine (A, code PDB : 3C22) et de DC-SIGN (B, code PDB : 2IT5).

### 13.3 Etude structurale de la Langérine avec les sucres de type GlcNac, GlcNac6S et GlcNS6S

Pour cette étude structurale, les cristaux utilisés pour le trempage avec les sucres ont été obtenus à partir du domaine ECD de la Langérine. En effet, due à la difficulté d'obtenir des quantités suffisantes de Lg-S-CRD, je me suis attelé à essayer de cristalliser le domaine ECD de la Langérine, mais cela n'a rien donné. En parallèle, j'ai réalisé des plaques de cristallisation dans les conditions proches de celles obtenues pour le domaine CRD. Après 2 à 3 mois, j'ai observé des cristaux avec un faciès identique à celui des cristaux obtenus avec la Lg-S-CRD. Nous avons donc réalisé une expérience de diffraction afin d'identifier si ces cristaux étaient composés, en totalité ou en partie, du domaine ECD. L'analyse des données cristallographiques conduit à un groupe d'espace et des paramètres de maille identiques à ceux de la Lg-S-CRD. Ceci suggère donc une protéolyse spontanée, dans la goutte, de la Langérine ECD pour former des cristaux de Langérine CRD. Pour confirmer cette hypothèse, j'ai récupéré une dizaine de cristaux que j'ai lavé dans la solution du puit correspondant et que j'ai ensuite dissout dans de l'eau afin de réaliser un séquençage N-Terminal (Annexe). Le résultat du séquençage N-Ter indique en effet une protéolyse au niveau des résidus du début du domaine CRD (résidus 190 de la séquence de la Langérine entière).

### 13.3. ETUDE STRUCTURALE DE LA LANGÉRINE AVEC LES SUCRES DE TYPE GLCNAC, GLCNAC6S ET GLCNS6S

Les informations confirmant que ces cristaux sont composés de Lg-CRD, et compte tenu du nombre important de cristaux obtenus, je les ai donc utilisé pour les trempages avec les différents monosaccharides étudiés en SPR. Les statistiques des données obtenues sur ces cristaux et celles d'affinement sont résumées dans le tableau 13.1.

Données/cristal	GlcNac	GlcNac6S	GlcNS6S
<b>Intégration</b>			
Groupe d'espace	P4 <sub>2</sub>	P4 <sub>2</sub>	P4 <sub>2</sub>
Paramètres de maille	a=b=80,241 Å; c=91,345 Å	a=b=79,753 Å; c=91,143 Å	a=b=79,483 Å; c=90,71 Å
Résolution (Å)	50-1,90 (2,00-1,90)	50-1,51 (1,60-1,51)	50-1,85 (1,96-1,84)
Nombre de réflexions uniques	44984 (7107)	86128 (13576)	48373 (7677)
Rmerge	0,05 (0,337)	0,051 (0,291)	0,047 (0,3)
I/sigma	13,6 (2,3)	13,9 (2,76)	18,98 (4,63)
Complétude (%)	98,5 (97,0)	97,4 (95,1)	99,5 (98,1)
<b>Phasage par remplacement moléculaire</b>			
TFZ	43,5	40	39,9
LLG	749	1379	1488
<b>Affinement</b>			
loi de macle	h, -k, -l		
Rwork (%)	14,41	19,72	17,39
Rfree (%)	17,35	22,69	21,08
Nombre réflexions utilisées	45058	86453	48410
Nombre atome de la protéine	4283	4302	4350
Nombre molécule d'eau	474	686	510
Nombre atome d'ion	9	9	11
Nombre atome saccharide	60	72	76
B moyen de la protéine (Å <sup>2</sup> )	26,67	26,51	29,6
B moyen des molécules d'eau (Å <sup>2</sup> )	32,42	34,31	35,2
B moyen ions autre calcium (Å <sup>2</sup> )	27,19	36,8	31,7
B moyen calcium (Å <sup>2</sup> )	18,98	28,8	30,05
B moyen saccharides (Å <sup>2</sup> )	40,2	32,3	43,7
% résidus dans la zone très favorable du Ramachandran	96,9	98,6	98,4
% résidus dans la zone favorable du Ramachandran	3,1	1,4	1,6
% résidus dans la zone interdite du Ramachandran	0	0	0
RMSD bond	0,006	0,007	0,007
RMSD angle	0,979	1,115	1,087

Tableau 13.1 – Statistiques d'intégration, de phasage par remplacement moléculaire et d'affinement des différentes structures.

Les structures résolues montrent une densité électronique claire autour des monosaccharides (Figure 13.7). Le taux d'occupation du GlcNac est de 0,8 contrairement au GlcNac6S et GlcNS6S pour lequel il est de 1. Ceci reflète probablement les différences d'affinité du GlcNac et des GlcNac6S et GlcNS6S. Le cristal ayant servi au trempage avec le GlcNac est maclé, ce qui par conséquent affecte les données.



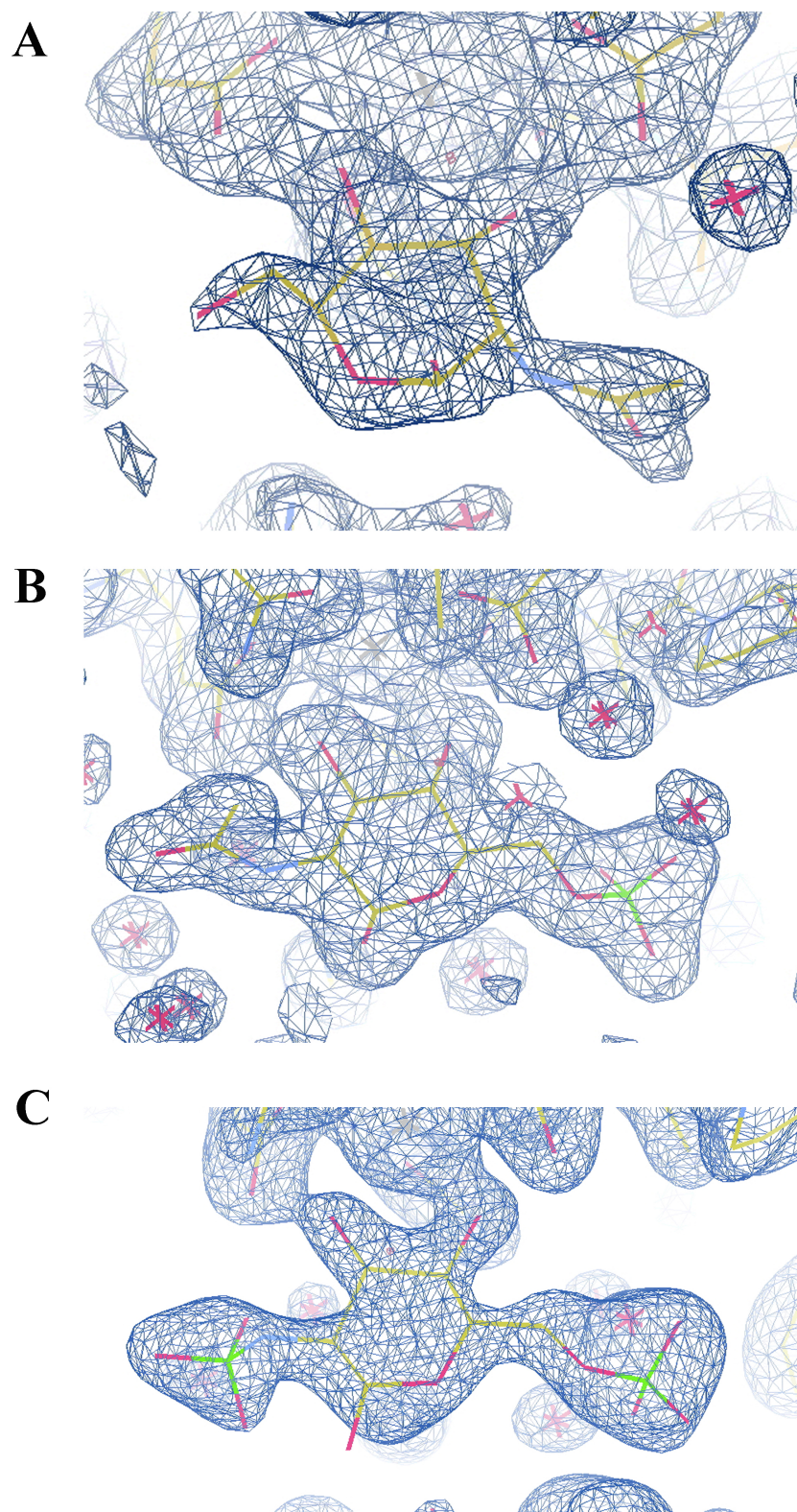


Figure 13.7 – Carte de densité électronique  $2F_O - F_C$  au niveau des saccharides (contour : 1 sigma). A : GlcNac. B : GlcNac6S. C : GlcNS6S.

### 13.3. ETUDE STRUCTURALE DE LA LANGÉRINE AVEC LES SUCRES DE TYPE GLCNAC, GLCNAC6S ET GLCNS6S

---

Les différentes structures révèlent une différence dans l'orientation des saccharides (Figure 13.8). En effet, le GlcNac est orienté de manière à ce que le groupement N-acétyl pointe en direction de la lysine 313 alors que dans les structures avec le GlcNac6S et le GlcNS6S, le groupement sulfate en position C6 s'oriente vers cette lysine. La présence du groupement sulfate en position C6 induit une rotation de 180° du sucre par rapport au calcium. C'est donc le groupement sulfate qui impose l'orientation des saccharides et qui est probablement à l'origine de l'amélioration de l'affinité par l'établissement d'une liaison électrostatique.

Outre l'interaction impliquant le calcium, cela permet une interaction directe entre la lysine 299 et l'hydroxyle anomérique (Figure 13.8 D-F) mais cette lysine, dans le cas du GlcNac, est en interaction avec l'amine. Cette interaction n'explique donc pas le gain d'affinité. Afin que les sucres sulfatés puissent interagir, la lysine 313 effectue un léger mouvement pour permettre au groupement sulfate de se placer correctement et éviter ainsi un encombrement stérique (1,8 Å, en prenant la structure avec le GlcNac, contre 2,1 Å dans la structure avec le GlcNac6S). Par conséquent, le réseau mis en place dans l'interaction est légèrement plus important dans le cas des saccharides sulfatés avec surtout une nouvelle liaison électrostatique entre la lysine 313 et le groupement sulfate en C6.

Le deuxième groupement sulfate du GlcNS6S n'intervient pas dans le mode de liaison (Figure 13.8 E-F), ceci explique le fait qu'il n'y a aucune différence significative d'affinité pour la Langérine entre le GlcNac6S et le GlcNS6S (voir Figure 13.6).

L'orientation des monosaccharides sulfatés permet aussi d'expliquer la diminution de l'affinité pour le trisaccharide Tri1a. Dans celui-ci, la liaison osidique reliant le GlcNac6S à l'IdoA (acide iduronique) est effectuée par le groupement hydroxyle anomérique du GlcNS6S en configuration alpha. Or dans cette configuration, qui correspond à l'orientation des structures avec le GlcNac6S et GlcNS6S, cet hydroxyle, positionné en alpha, est en interaction directe avec la lysine 299. Par conséquent, l'IdoA serait directement orienté vers le squelette de la protéine et serait donc en encombrement stérique avec celle-ci. Pour éviter ce clash stérique, il est probable que le Tri1a s'oriente comme le GlcNac et ne peut présenter son sulfate en position C6 vers la lysine 313. Dans ce modèle, le Tri1a ne bénéficierait donc pas de son sulfate pour l'interaction avec la protéine qui se ferait principalement *via* les hydroxyles en C3 et C4 expliquant ainsi l'affinité similaire entre le Tri1a et le GlcNac. Un jeu de données, d'un cristal trempé en présence de Tri1a, n'a pas été encore traité et permettrait probablement d'observer l'orientation du Tri1a.

Finalement, quelque soit l'orientation du sucre, le réseau d'interaction mis en place avec les hydroxyles en C3 et C4 est le même, il implique des interactions de ces hydroxyles avec le calcium ainsi que des interactions entre le glutamate 285 et l'hydroxyle en C4 pour le GlcNac et C3 pour les sucres sulfaté ainsi qu'entre le glutamate 293 et l'hydroxyle en



C3 pour le GlcNac et en C4 pour les deux autres sucres. Donc, l'orientation du sucre n'influe pas sur la capacité de ces groupes hydroxyles à chélater le calcium et interagir avec la protéine.

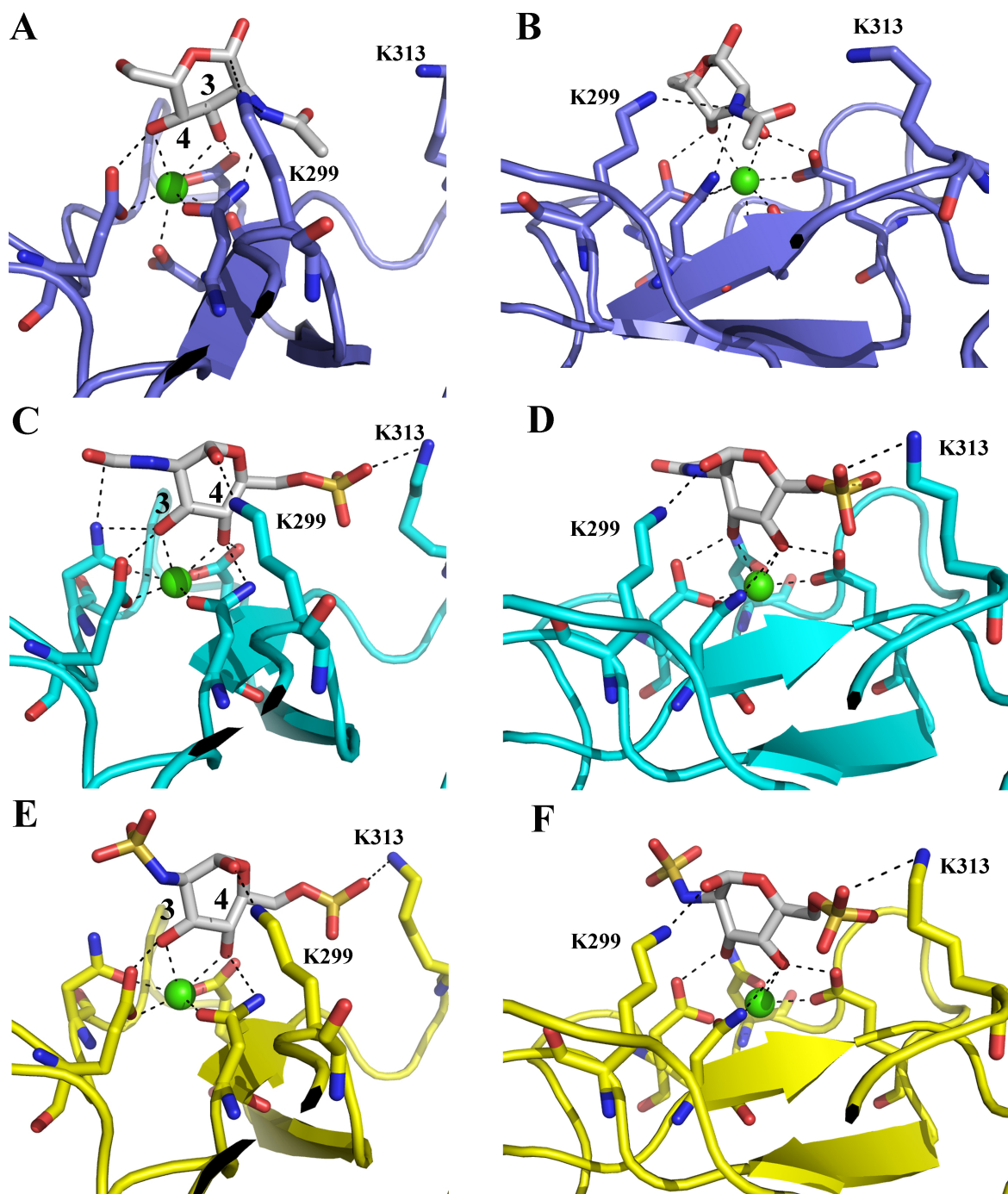


Figure 13.8 – Structure cristallographique mettant en évidence l'orientation des saccharides au niveau du site calcium. A-B : GlcNac. C-D : GlcNac6S. E-F : GlcNS6S. Les pointillés représentent le réseau de liaisons impliquées dans l'interaction (distance entre 2,5 et 3,5 Å).

## 13.4 Etude par RMN de l'interaction avec des trisaccharides sulfatés

La méthode utilisée par RMN est le transfert de saturation des protons de la protéine aux protons du ligand (Figure 13.9). Cette méthode permet d'avoir des données sur le ligand et non sur la protéine et permet ainsi une interprétation du signal plus facile. En effet, cette méthode est basée sur la différence de deux spectres du proton à une dimension. Le premier spectre est celui de tous les protons des molécules en solution et le second est le spectre du proton après la saturation sélective des protons de la protéine. Le premier spectre est donc effectué avec le ligand seul en solution, puis le second est réalisé en présence de protéine et de ligand en saturant sélectivement les protons de la protéine. Ainsi, le transfert de saturation des protons de la protéine vers les protons du ligand va affecter l'intensité des signaux. Plus le transfert est important, plus les signaux des protons du ligand impliqués dans le transfert vont diminuer. C'est pour cela que la soustraction des deux spectres permet de mettre en évidence les protons du ligand proche de la protéine et de quantifier le transfert de saturation. Plus le transfert de saturation est important, plus les protons sont proches de la protéine. De ce type d'expérience, nous pouvons identifier les portions du ligand importantes pour l'interaction avec la protéine.

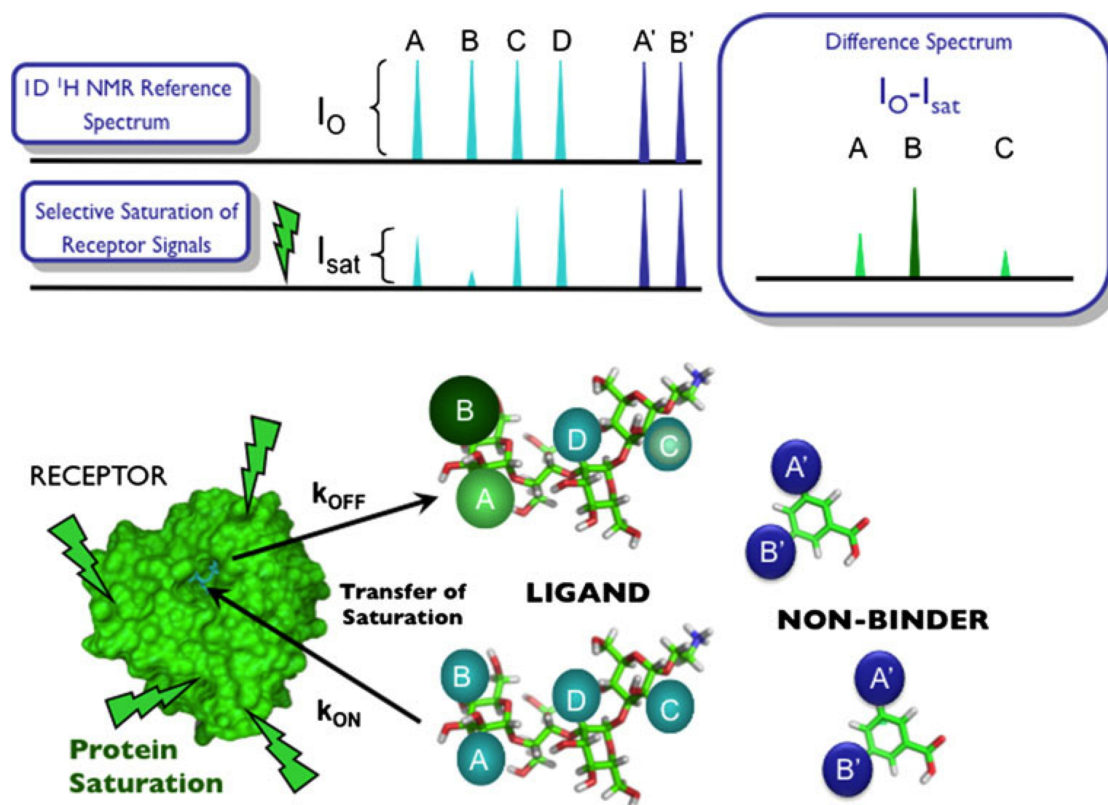


Figure 13.9 – Schéma basique d'une expérience de STD.

L'objectif de départ, de nouveau, était d'étudier le site d'interaction avec les glycosaminoglycanes. La RMN nous permettait de nous affranchir des problèmes de cristallisation de la Langérine ECD. Mais, les problèmes principaux des expériences de RMN STD avec des molécules fortement chargées sont, tout d'abord, une interaction forte qui ne permet pas d'observer un signal de transfert de saturation, puis que la taille des groupements sulfates induit des distances entre les protons de la protéine et ceux du ligand importantes pour permettre un transfert de saturation. Pour ces deux raisons, nous avons commencé avec des trisaccharides car ils possèdent moins de groupement sulfate qu'avec un fragment important d'héparine, et une affinité suffisamment « faible » pour observer un signal STD.

Ces expériences de RMN sont réalisées en collaboration avec Pedro M. Nieto et Jesus Angulo (IIQ, CSIC, Séville). Outre l'accès à l'outil RMN, ce groupe a une expertise dans la synthèse d'oligosaccharide sulfaté. Notamment, ils ont développé la synthèse de deux séries de trisaccharides présents dans la chaîne d'héparane sulfate (Figure 13.10).

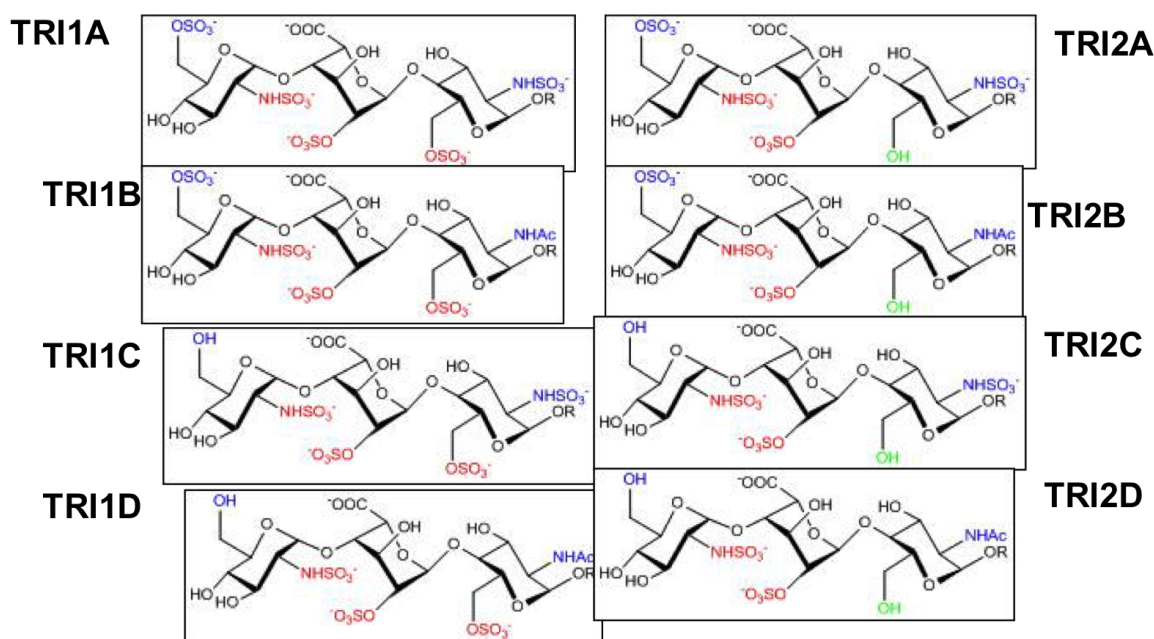


Figure 13.10 – Représentation des deux séries de trisaccharides de synthèse produites par nos collaborateurs.

La différence entre les séries 1 et 2 est la présence d'un groupement sulfate en position C6 du glucosamine réducteur. La différence entre les trisaccharides A et B est la présence du groupement sulfate en N2 sur le glucosamine réducteur. Enfin, les trisaccharides de série C et D se différencient respectivement des séries A et B par la perte du sulfate en position C6 du glucosamine non réducteur. Ils disposent ainsi d'une librairie de 8 trisaccharides correspondant à des motifs saccharidiques de type héparine. Dans le cadre de ce travail, nous avons pu utiliser à ce jour les trisaccharides Tri1a et Tri2b. La différence

entre ces deux trisaccharides Tri1a et Tri2b est la présence de deux groupements sulfate supplémentaires en position N2 et C6 du glucosamine réducteur pour le Tri1a (Figure 13.11).

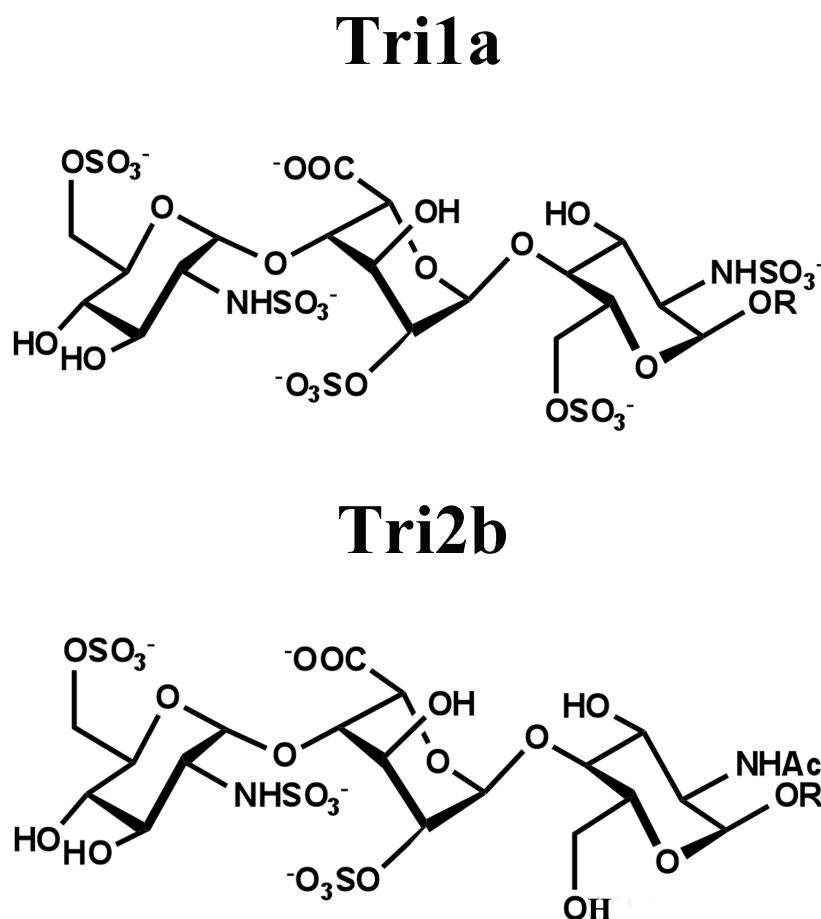


Figure 13.11 – Structure du trisaccharide Tri1a et Tri1b.

Les premiers résultats ont montré que l'interaction de la Langérine avec ces trisaccharides est dépendante du site calcium car en présence d'EDTA cette interaction n'est plus observable (Figure 13.12 B-D). Ainsi même en présence d'EDTA, qui permet dans le cas d'une interaction dans les deux sites simultanément de masquer l'interaction dans le site calcique, un trisaccharide ne semble pas capable d'interagir dans le site héparine. En présence de calcium, on est capable d'observer un signal et d'attribuer quels hydrogènes sont en interaction avec la protéine ou à proximité. Comme pour la cristallographie, ces composés nous ont orienté vers la caractérisation du mode d'interaction de sucre sulfatés dans le site calcique de la Langérine (chronologiquement les expériences de SPR ont été réalisées en dernier).

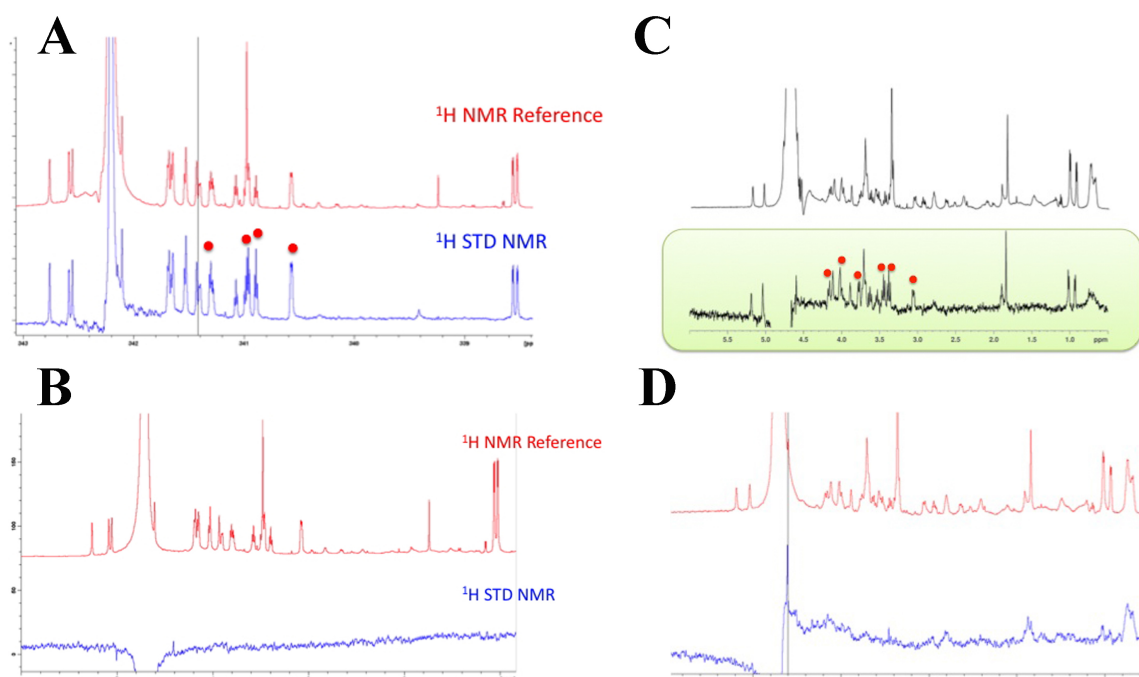


Figure 13.12 – Spectre 1D du proton de l'expérience STD avec le Tri1a en calcium (A) et en EDTA (B) et avec le Tri1b en calcium (C) et en EDTA (D). Les points rouges mettent en évidence les protons pour lesquels il y a eu échange de saturation. C : le spectre du haut est le spectre de référence et celui du bas est le spectre après soustraction avec le spectre STD. D : le spectre du haut est le spectre de référence et celui du bas est le spectre STD.

En effet, lorsque l'on regarde la structure de ces trisaccharides, on observe à l'extrémité non réductrice de l'oligosaccharide la présence d'un GlcNac6S possédant les groupes hydroxyles en C3 et C4 libres. Or, le groupe hydroxyle en C4 est normalement utilisé dans la liaison osidique de la chaîne d'héparane sulfate. Il semble qu'il est suffisant pour ce type de sucre de posséder les hydroxyles en C3 et C4 libres pour rendre l'interaction dépendante du calcium. Probablement est ce dû aussi au fait qu'un trisaccharide ne possède pas une affinité suffisante dans le site héparine indépendant du calcium pour être observé par cette technique et de conforter l'idée d'un site d'interaction étendu pour l'héparine. En effet, on a pu voir dans le chapitre précédent qu'un hexasaccharide était une taille limite pour observer une inhibition par SPR.

Cette expérience de STD permet de quantifier la proximité de chaque hydrogène du ligand par l'efficacité du transfert de saturation. En effet, on normalise l'intensité des pics en fonction du plus intense qui est arbitrairement évalué comme 100% du transfert de saturation. Ceci permet d'évaluer ensuite l'importance de chaque groupement hydroxyle (Figure 13.13). On remarque évidemment une forte interaction du saccharide de l'extrémité non réduite avec notamment les hydroxyles en C3 et C4. On peut observer aussi des



### 13.4. ETUDE PAR RMN DE L'INTERACTION AVEC DES TRISACCHARIDES SULFATÉS

interactions supplémentaires par les hydroxyles des saccharides suivants ce qui indique une certaine proximité avec la protéine.

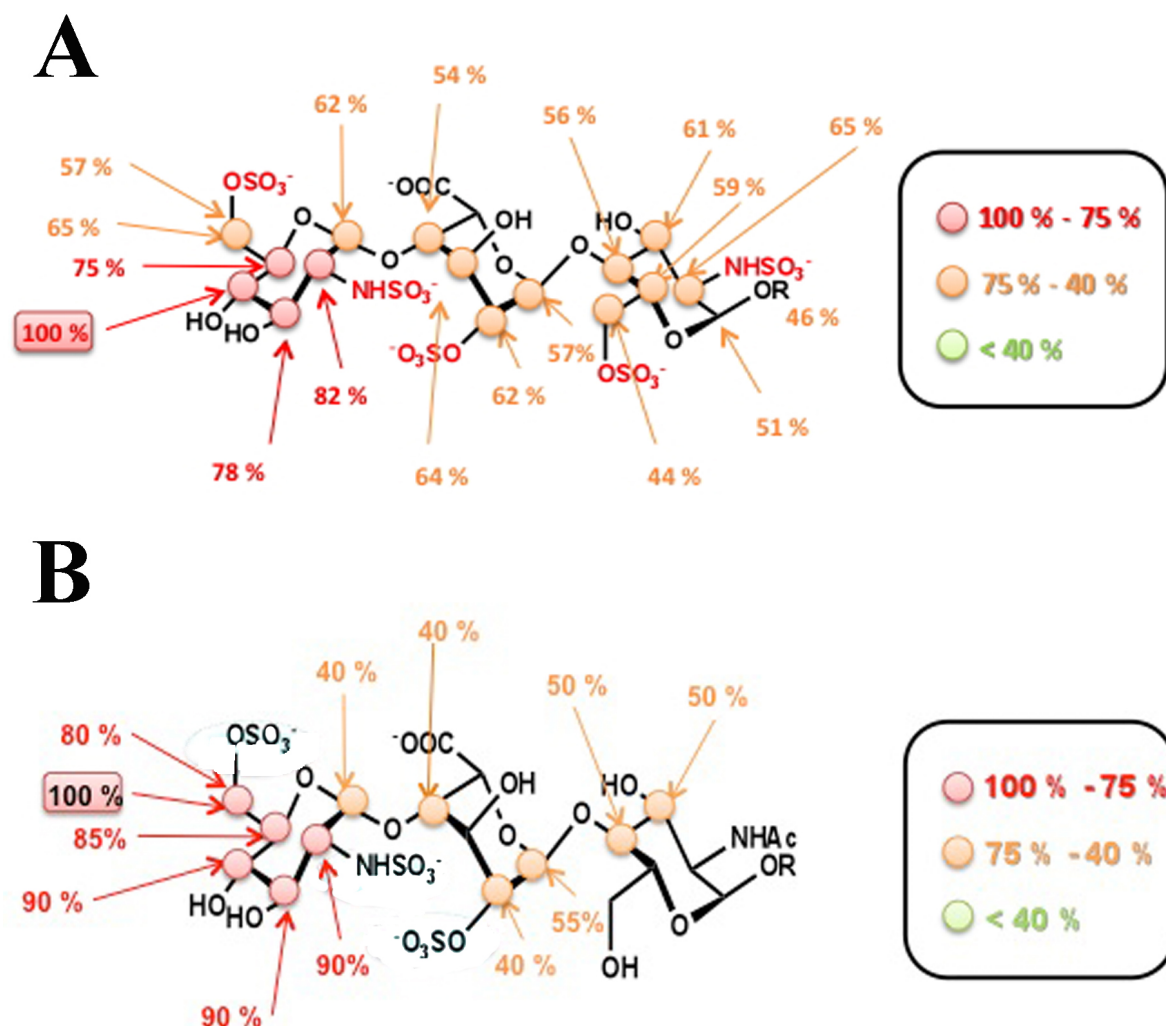


Figure 13.13 – Quantification du transfert de saturation des protons sur le Trila (A) et le Tri2b (B).

On observe des différences de transfert d'aimantation aux protons aux extrémités réduite et non réduite du Tri2B par rapport au Trila. Il semble que dans le cas du Tri2b, le glucosamine non réduit est plus proche de la protéine que pour le Trila par une différence significative du transfert de saturation aux hydrogènes en C6 notamment. Par contre, sur le reste de la chaîne, le transfert de saturation aux hydrogènes est moins prononcé sur les deux autres sucres. Ces différences de transfert de saturation suggèrent que les groupements sulfates, de l'acide iduronique et du glucosamine à l'extrémité réductrice, rapprochent l'ensemble de la chaîne vers la protéine en déstabilisant peut être la partie en interaction avec le calcium.

Enfin, nos collaborateurs ont montré dans une expérience préliminaire récente, qu'avec un hexasaccharide, que l'interaction était observable en tampon EDTA indiquant une interaction indépendante du site calcium. Cet hexasaccharide atteint probablement la taille limite pour observer une interaction dans le site héparine. De plus, il possède à son extrémité non réductrice un groupement sulfate en C4 bloquant celui-ci pour toute interaction dans le site calcique de la Langérine (Figure 13.14). Cet hexasaccharide est donc un outil de premier plan pour la caractérisation par RMN du site héparine.

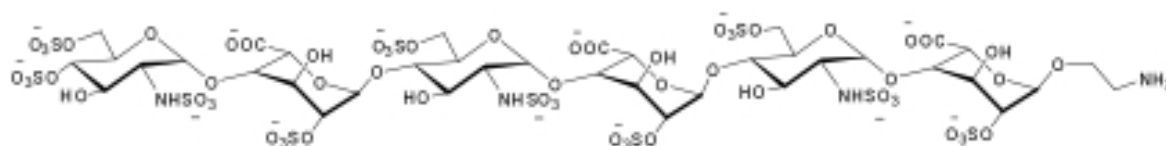


Figure 13.14 – Structure de l'hexasaccharide utilisé.

## 13.5 Conclusion

Une structure publiée en 2011 montre une double orientation possible du mannose en interaction avec la Langérine (Figure 13.15) [40]. La différence principale entre ces deux positions du mannose est l'orientation du cycle qui permet de présenter le groupement hydroxyle anomérique de manière différente (Figure 13.15, B). Dans cette nouvelle orientation, l'hydroxyle anomérique pointe vers la lysine 299. Cette présentation du groupement hydroxyle anomérique permet l'établissement d'une liaison hydrogène avec cette lysine. Dans la deuxième orientation la lysine 299 pointe vers l'hydroxyle en C2 mais la distance est dans la limite du raisonnable (Figure 13.15, A). Ces deux orientations sont possible pour un monosaccharide mais dans le cas d'oligosaccharide on s'attend à obtenir l'orientation de la figure 13.15 A lorsque que la liaison osidique, qui implique l'hydroxyle anomérique, est en alpha ou en bêta et l'orientation de la figure 13.15 B lorsque celle-ci est seulement en bêta.

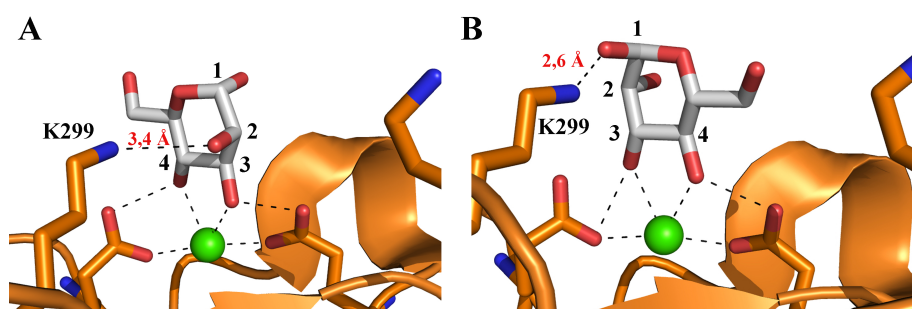


Figure 13.15 – Structure cristallographique du CRD de la Langérine en présence de mannose (ID : 3P5F).

La présence du groupement N-acétyle du glucosamine impose une orientation où l'hydroxyle anomérique n'interagit pas avec la lysine 299. Par contre en solution, la double orientation doit exister comme avec le mannose (Figure 13.15). Ceci pourra être confirmé ultérieurement par des expériences de STD avec ce monosaccharide.

Dans le cas des monosaccharides sulfatés, nous avons observé une amélioration significative de l'affinité par SPR qui est expliquée, dans les structures cristallines, par l'établissement de liaisons supplémentaires. En effet, le groupement sulfate 6 impose l'orientation du cycle saccharidique de façon à stabiliser la liaison entre l'hydroxyle anomérique et la lysine 299 mais aussi de pointer le groupement sulfate en C6 vers la lysine 313. Le second sulfate du GlcNS6S n'apporte aucune interaction supplémentaire expliquant la même affinité que pour le monosulfaté.

Pour les trisaccharides (Tri1a et Tri1b), nous avons observé par SPR une affinité similaire que pour le monosaccharide non sulfaté. Or, les données de STD montre un transfert de saturation pour les hydrogènes du deuxième et troisième saccharides. Ce transfert de saturation observé devrait expliquer une amélioration de l'affinité par l'établissement de liaisons supplémentaires avec la protéine. Un jeu de données cristallographique a été obtenu pour le CRD de la Langérine en présence du Tri1a. Les composés utilisés, pour le moment, dans les expériences de SPR et de RMN diffèrent au niveau du saccharide de l'extrémité réduite. Le Tri1a possède deux groupements sulfates sur ce dernier saccharide que ne possède pas le Tri2b. Ces sulfates jouent peut être un rôle d'attraction dans l'interaction avec la protéine qui n'existe peut être pas avec le Tri1b d'où un meilleur transfert de saturation au niveau de l'extrémité réductrice du Tri1a. L'expérience de SPR n'a pas pu être réalisée avec le Tri1b car nos collaborateurs ne possédaient pas suffisamment de ce trisaccharide.

Pour conclure, la présence des résidus lysines 299 et 313 de la Langérine procure la spécificité de reconnaissance pour les saccharides sulfatés comparé à DC-SIGN. Cette spécificité est probablement importante physiologiquement mais les saccharides sulfatés physiologiques sont principalement présents dans les chaînes de glycosaminoglycanes qui, comme nous avons pu le voir précédemment, n'interagissent pas avec la Langérine dans le site calcium. Donc, la pertinence de ce type d'interaction au niveau physiologique est encore à démontrer. On peut suggérer que l'interaction avec des oligosaccharides sulfatés est favorable si celui-ci possède un glucosamine sulfaté en C6 avec les hydroxyles en C3 et C4 libres et qu la liaison osidique est réalisée en bêta sur l'hydroxyle anomérique du sucre en interaction avec le calcium, car si-non le sulfate n'apporte aucun bénéfice s'il est orienté comme un GlcNac. En effet, si la liaison est en alpha, alors le sucre ne peut pas s'orienter comme le GlcNac6S sans être encombré par la protéine.

Des études ont montré l'importance de GlcNac6S dans la communication cellule-cellule



comme, par exemple, la modification des glycosylations de CD44 par la présence de  $\text{TNF-}\alpha$  qui induit une augmentation de LacNac6S ( $\text{Gal}\beta 1\text{-4GlcNac6S}$ )[27]. Ce même oligosaccharide (LacNac6S) est un épitope de cellule dendritique inflammatoire [88] mais la structure de LacNac6S empêche une liaison dans le site calcium car ce sucre est composé normalement d'un galactose lié en  $\beta 1\text{-4}$  à un glucosamine sulfaté en C6. Or Tatenos et al affirme observer une interaction de la Langerin avec le 6-sulfo-LacNac et que cette interaction serait dépendante de la lysine 313 mais différente de l'interaction avec seulement le 6-sulfo-glucosamine [99]. En effet, ils ont montré que c'était l'épitope Galactose sulfaté en position 6 qui permettait l'interaction avec la Langérine du LacNac sulfaté. On retrouve donc le même épitope Gal6S que précédemment décrit dans l'étude de Feinberg [40]. Enfin en comparaison avec DC-SIGN qui a été tout d'abord identifiée comme une protéine d'adhésion cellulaire, la Langérine, par cette spécificité de reconnaissance, pourrait avoir un rôle dans l'adhésion cellulaire outre son rôle de PRR et dans la formation de BG. Des études cellulaires sont nécessaires pour élucider le rôle de la Langérine pour ce type de saccharide.

---

## Chapitre 14

# Développement d'inhibiteurs de DC-SIGN

Tout d'abord, ce travail de développement est effectué au sein d'un projet de recherche européen. Il consiste à générer des ligands spécifiques de DC-SIGN. L'objectif est d'obtenir des ligands candidats pour empêcher l'infection par le VIH. La spécificité pour DC-SIGN est donc importante car la Langérine est impliquée dans un mécanisme empêchant l'infection par ce virus.

Ensuite, ce travail a conduit précédemment au dépôt d'un brevet pour la molécule « pseudotrimannose » qui est utilisée dans le premier article de ce chapitre. Cette article prouve la stratégie envisagée par ce réseau européen. En effet, ce travail montre qu'un ligand composé de mannose et « pseudo-mannose » présenté en dendron permet d'inhiber une infection du virus au niveau des muqueuses génitales. Le dendron est un système de présentation multivalent de ligands qui permet d'augmenter l'affinité pour les récepteurs, DC-SIGN. L'étude de SPR réalisée avec les ligands non présentés en dendron montre une meilleure affinité pour DC-SIGN par rapport à la Langérine. Mais dans le contexte d'inhibiteur strictement spécifique de DC-SIGN, le développement des ligands à base « mannose » est encore à améliorer comme, par exemple, par l'ajout de charge positive afin de créer une répulsion avec les lysines 299 et 313 de la Langérine.

Enfin, le deuxième article montre la forte différence de reconnaissance de DC-SIGN et de la Langérine pour des composés à base fucose. En effet, les ligands étudiés dans ce travail inhibent beaucoup plus, pour une même concentration en ligand, DC-SIGN que la Langérine. La stratégie menée avec des inhibiteurs à base « fucose » est potentiellement préférable à celle menée avec une base mannose pour le développement d'inhibiteurs strictements spécifiques de DC-SIGN.

Dans ces différents travaux, ma contribution s'est limitée à la production et la purification de la Langérine ainsi qu'à un support dans l'utilisation de la Langérine pour

---

les expériences de SPR. Ces expériences de SPR ont été réalisées par une doctorante de l'équipe, Ieva Sutkeviciute, affectée exclusivement à ce projet. Du fait de ma contribution, même modeste, j'ai été associé à ces deux articles qui sont ici pour mémoire.

# A glycomimetic compound inhibits DC-SIGN-mediated HIV infection in cellular and cervical explant models

Angela Berzi<sup>a</sup>, José J. Reina<sup>d</sup>, Roberta Ottria<sup>d</sup>, Ieva Sutkeviciute<sup>e,f,g</sup>,  
Patrizio Antonazzo<sup>i</sup>, Macarena Sanchez-Navarro<sup>j</sup>, Eric Chabrol<sup>e,f,g</sup>,  
Mara Biasin<sup>a</sup>, Daria Trabattoni<sup>a</sup>, Irene Cetin<sup>i</sup>, Javier Rojo<sup>j</sup>,  
Franck Fieschi<sup>e,f,g,h</sup>, Anna Bernardi<sup>d,k</sup> and Mario Clerici<sup>b,c</sup>

**Objective:** Dendritic cell-specific intercellular adhesion molecule (ICAM)-3 grabbing nonintegrin (DC-SIGN) participates in the initial stages of sexually transmitted HIV-1 infection by recognizing highly mannosylated structures presented in multiple copies on HIV-1 gp120 and promoting virus dissemination. Inhibition of HIV interaction with DC-SIGN thus represents a potential therapeutic approach for viral entry inhibition at the mucosal level.

**Design:** Herein we evaluate the efficacy in inhibiting HIV-1 infection and the potential toxicity of a multimeric glycomimetic DC-SIGN ligand (Dendron 12).

**Methods:** The ability of Dendron 12 to block HIV-1 infection was assessed in cellular and human cervical explant models. Selectivity of Dendron 12 towards DC-SIGN and langerin was evaluated by surface plasmon resonance studies.  $\beta$  chemokine production following stimulation with Dendron 12 was also analyzed. Toxicity of the compound was evaluated in cellular and tissue models.

**Results:** Dendron 12 averted HIV-1 *trans* infection of CD4<sup>+</sup> T lymphocytes in presence of elevated viral loads and prevented HIV-1 infection of human cervical tissues, under conditions mimicking compromised epithelial integrity, by multiple clades of R5 and X4 tropic viruses. Treatment with Dendron 12 did not interfere with the activity of langerin and also significantly elicited the production of the  $\beta$  chemokines MIP-1 $\alpha$ , MIP-1 $\beta$  and RANTES.

**Conclusion:** Dendron 12 thus inhibits HIV-1 infection by competition with binding of HIV to DC-SIGN and stimulation of  $\beta$ -chemokine production. Dendron 12 represents a promising lead compound for the development of anti-HIV topical microbicides.

© 2012 Wolters Kluwer Health | Lippincott Williams & Wilkins

*AIDS* 2012, **26**:127–137

**Keywords:** cervical explants, DC-SIGN, glycomimetic drugs, HIV, topical microbicides

---

<sup>a</sup>Chair of Immunology, Department of Clinical Sciences L. Sacco, <sup>b</sup>Chair of Immunology, Department of Biomedical Sciences and Technologies, University of Milan, Segrate, <sup>c</sup>Don C. Gnocchi Foundation IRCSS, <sup>d</sup>Department of Organic and Industrial Chemistry, University of Milan, Milan Italy, <sup>e</sup>Institut de Biologie Structurale, Université Joseph Fourier, <sup>f</sup>CEA, DSV, <sup>g</sup>CNRS, UMR 5075, Grenoble, <sup>h</sup>Institut Universitaire de France, Paris, France, <sup>i</sup>Unit of Obstetrics and Gynaecology, Department of Clinical Sciences L. Sacco, University of Milan, Milan, Italy, <sup>j</sup>Instituto de Investigaciones Químicas, CSIC - University of Sevilla, Sevilla, Spain, and <sup>k</sup>CNR-ISTM, Institute of Molecular Science and Technologies, Milan, Italy.

Correspondence to Mario Clerici, Chair of Immunology, Department of Biomedical Sciences and Technologies, University of Milan, Via Fratelli Cervi 93, 20090 Milan, Italy.

Tel: +39 025030412; fax: +39 025030414; e-mail: mario.clerici@unimi.it

Received: 7 September 2011; revised: 5 October 2011; accepted: 12 October 2011.

DOI:10.1097/QAD.0b013e32834e1567

## Introduction

HIV remains one of the leading causes of mortality and morbidity [1]. As the vast majority of HIV-1 infections occur via sexual transmission through mucosal surfaces, the development of vaginal and rectal topical microbicides represents a promising approach to prevent sexually transmitted HIV-1 infection.

Dendritic cell-specific intercellular adhesion molecule-3 grabbing nonintegrin (DC-SIGN) is involved in the initial step of HIV-1 sexually transmitted infection and it may be considered a promising therapeutic target [2,3].

Myeloid immature dendritic cells located in mucosal tissue of vagina, cervix and rectum express DC-SIGN and are among the first cells to encounter sexually transmitted HIV [4–6]. Upon HIV-1 binding to DC-SIGN, the virus escapes (at least partially) degradation into lysosomes. Rather, it is internalized into endosomes and multivesicular bodies, in which it is protected from degradation and retained in a high infective state [7–9]. Dendritic cells transmit the virus in *trans* to adjacent CD4<sup>+</sup> T lymphocytes in genital mucosae or, after migration to lymphoid tissue, to CD4<sup>+</sup> T lymphocytes resident in lymphoid tissue, promoting HIV-1 dissemination [10,11]. DC-SIGN, facilitating HIV-1 interaction with CD4 and co-receptors, also enhances dendritic cell infection in *cis*, that results in long-term transmission of HIV [12,13].

In addition, binding of HIV-1 to DC-SIGN activates signaling pathways that modulate Toll-like receptor signaling, inducing immunosuppressive responses and triggering HIV replication and transmission [14–17].

DC-SIGN specifically recognizes the high mannose glycan (Man<sub>9</sub>), presented in multiple copies on HIV-1 gp120 [8,18]. Fragments of Man<sub>9</sub> terminated by a di-mannoside or a tri-mannoside bind to DC-SIGN almost as efficiently as the entire Man<sub>9</sub> [19]. Chemically synthesized analogs of (Man)<sub>9</sub> terminal di-saccharides and tri-saccharides are more resistant to hydrolysis by glycosydases than the corresponding natural oligosaccharides and interact efficiently with DC-SIGN [20]. These analogs can be linked to tetravalent (dendrons) scaffolds to obtain compounds endowed with stronger binding affinities to DC-SIGN [21,22]. Such compounds can compete with binding of HIV gp120 to DC-SIGN and are suitable for the development of new anti-HIV microbicides.

We recently demonstrated that a tetravalent dendron containing four copies of a linear pseudo-mannotriptide (Dendron 12) was able to inhibit HIV-1 *trans* infection of CD4<sup>+</sup> T lymphocytes [23].

Experimental models based on infection with HIV of explants taken from human uterine cervix, albeit with some limitations, allow a better approximation of the conditions *in vivo* compared to cellular models [24–27]. Thus, in this study a cervical explant model was exploited to assess the efficacy in inhibiting HIV-1 infection and the toxicity of Dendron 12, with the purpose of evaluating if the compound is a suitable candidate for the development as topical microbicide.

## Material and methods

### Cell culture

B-THP1 and B-THP1/DC-SIGN cells (contributed by Drs Li Wu and Vinet N. KevalRamani) were cultured as described [23].

### Virus

The following HIV-1 strains were used: BaL (contributed by Drs. S. Gartner, M. Popovic and R. Gallo; DU174 (Source: Dr L Morris); the R5 and X4 tropic primary isolates 8 g and DPMVF (provided by Professor Stefano Aquaro).

### Inhibition of HIV infection in *trans*

Human CD4<sup>+</sup> T lymphocytes were purified from peripheral blood of volunteer healthy donors following written consent, and activated as described [23]. Dendron 12, synthesized as described [23], was diluted to desired concentration into culture medium [Roswell Park Memorial Institute (RPMI) 160 with 20% fetal bovine serum (FBS), penicillin and streptomycin and L-glutamine, all from Euroclone, Sizzano, Italy]. B-THP1/DC-SIGN or B-THP1 cells (10<sup>6</sup> cells/ml) were preincubated with Dendron 12 (250 µmol/l), or culture medium alone for 30 min prior to exposure to BaL (virus titer ranging from 5 TCID<sub>50</sub> to 80 TCID<sub>50</sub>), in the continued presence of the inhibitor 3 h at 37°C. After extensive washing, B-THP1/DC-SIGN cells were co-cultured with activated human CD4<sup>+</sup> T cells as previously described [23].

### Human cervical explants infection

Cervical tissue was obtained, following written informed consent, from premenopausal women, HIV, hepatitis B virus (HBV) and hepatitis C virus (HCV) seronegative and without current genital infection, undergoing therapeutic hysterectomy at Unit of Obstetrics and Gynaecology of Sacco Hospital (Milan, Italy). The study was approved by the local ethic committee and conducted in compliance with international guidelines and local laws. Endocervical 3 mm × 3 mm explant biopsies comprised both epithelium and stromal tissue.

Within 1 h after obtaining tissues the explants were extensively washed. Then explants were either

immediately treated with the Dendron **12** and infected with R5 tropic strains or were prestimulated for 2 days in presence of IL-2 (R&D systems, Minneapolis, Minnesota, USA) and phytohaemagglutinin (PHA) (Sigma–Aldrich, Saint Louis, Missouri, USA) prior to X4 tropic DPMVF exposure.

Explants were pretreated 30 min at 37°C with different concentrations of Dendron **12**. Afterwards explants were exposed to HIV-1 BaL, DU174 (both  $2.6 \times 10^4$  TCID<sub>50</sub>), 8g or DPMVF (both  $10^4$  TCID<sub>50</sub>) in the continued presence of the compound 3 h at 37°C. Subsequently explants were washed with RPMI. Unstimulated explants were cultured in RPMI medium supplemented with 20% FBS, penicillin and streptomycin, L-glutamine (Euroclone) and gentamycin (Sigma–Aldrich) at 37°C and 5% CO<sub>2</sub>. Stimulated explants were cultured under the same conditions in presence of IL-2. Supernatants were collected 3 and 7 days post infection.

### p24 ELISA

p24 concentration in the supernatants was assayed by Alliance HIV-1 p24 Antigen kit (Perkin Elmer, Waltham, Massachusetts, USA). Plates were read using the IMark microplate reader equipped with Microplate Manager 6 software (Biorad, Segrate, Italy).

### Toxicity on peripheral blood mononuclear cells

Peripheral blood mononuclear cells (PBMC), purified from peripheral blood of healthy donors as described [23], were incubated with different concentrations of Dendron **12** for 3 or 7 days. The apoptosis was monitored evaluating the percentage of dead cells by staining with 7-AAD (Beckman Coulter, Fullerton, California, USA). Flow cytometric analyses were performed using a CYTOMICS FC-500 flow cytometer interfaced with CXP21 software (Beckman Coulter).

### 3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide assay

Toxicity of Dendron **12** was determined by a 3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide-based assay (Sigma–Aldrich). Viable explants reduce MTT to formazan crystals. Explants were cultured with increasing concentration of Dendron **12** diluted in medium culture for 3 and 7 days. The 50% of medium culture (containing the different concentration of Dendron **12**) was changed at day 3. After culturing, explants were washed and incubated in medium RPMI without phenol red and 10% FBS and MTT. Formazan was dissolved by MTT solubilization solution and formazan absorbance was measured at 595 nm.

### Differentiation and treatment of monocyte-derived dendritic cells

CD14<sup>+</sup> monocytes were separated from PBMCs using the CD14<sup>+</sup> microbeads (Miltenyi Biotech, Bergisch Gladbach, Germany) following manufacturer's instruction.

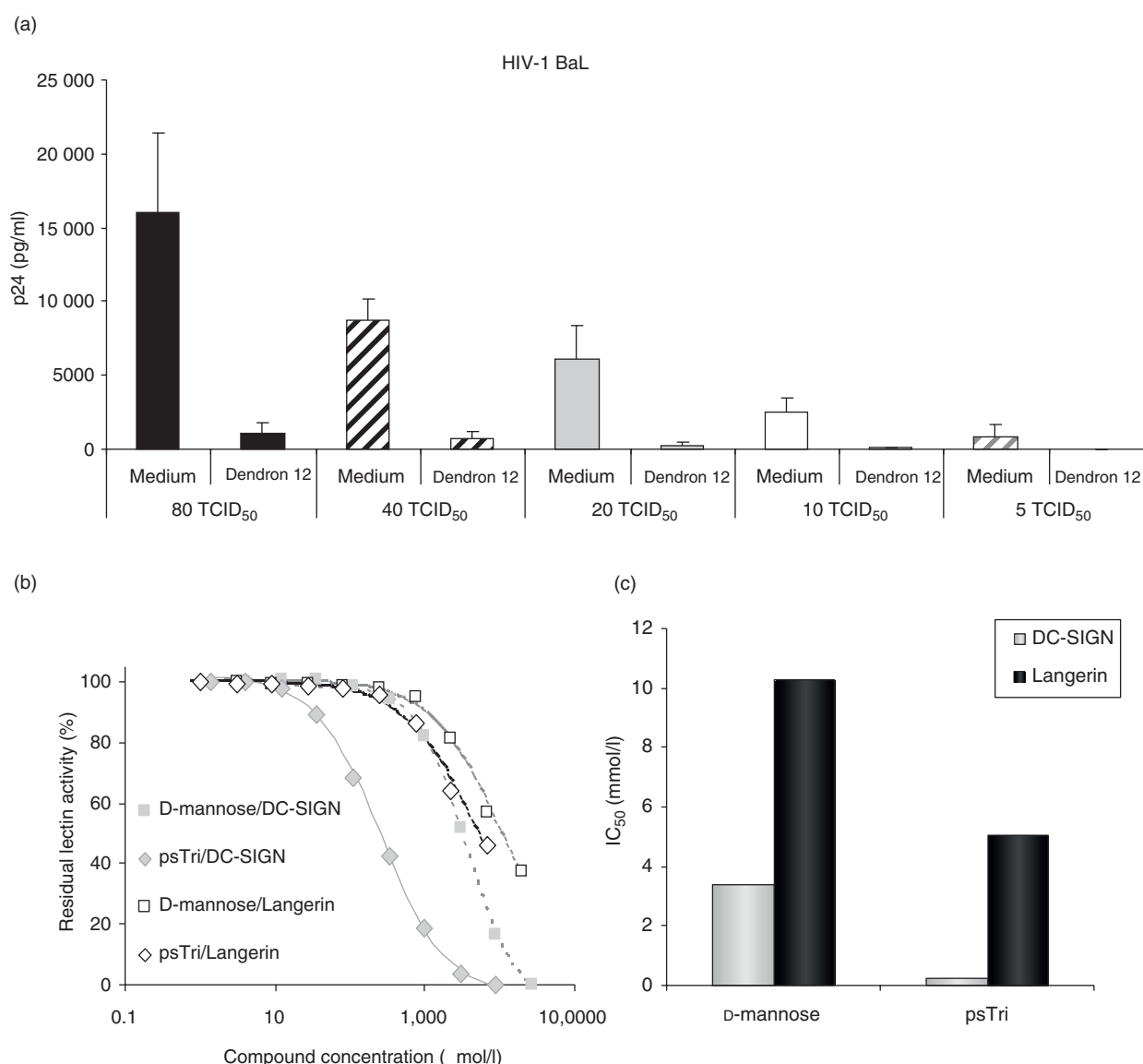
Monocytes were differentiated into iDCs by culturing them in presence of IL-4 (20 ng/ml) and granulocyte-macrophage colony-stimulating factor (20 ng/ml) (R&D Systems) for 6 days. DC-SIGN expression was checked by staining with anti human DC-SIGN-PE monoclonal antibody (clone AZND1, Beckman Coulter) and flow cytometric analysis. Monocyte-derived dendritic cells (MDDCs) were treated with Dendron **12**. Supernatants were harvested and RNA extracted from cells after 3, 24 and 72 h.  $\beta$  chemokine concentration in culture supernatants was evaluated using DuoSet kits (R&D Systems).

### RNA extraction and real-time PCR

RNA was extracted using the acid guanidium thiocyanate–phenol–chloroform method and purified from genomic DNA with RNase-free DNase (RQ1 DNase, Promega, Madison, Wisconsin, USA). RNA was reverse-transcribed using random hexamer primers and M-MLV reverse transcriptase (Clontech, Palo Alto, California, USA). cDNA quantification for macrophage inflammatory protein (MIP)-1 $\alpha$ , MIP-1 $\beta$ , RANTES (regulated upon activation, normal T-cell expressed, and secreted) and glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was performed by real-time PCR (DNA Engine Opticon 2; MJ Research, Ramsey, Minnesota, USA). Reactions were performed using a SYBR Green PCR mix (Finnzymes, Espoo, Finland). Results were expressed as  $\Delta\Delta C_t$  and presented as ratios between the target gene and the GAPDH housekeeping mRNA.

### Surface plasmon resonance analysis

Extracellular domain (ECD) of langerin (residue 68–328) and DC-SIGN (residue 66–404) was overexpressed and purified as described [28,29]. Surface plasmon resonance (SPR) experiments were performed on a Biacore 3000 using functionalized CM4 sensor chips and the corresponding reagents from Biacore. Two flow cells were activated as described [30]. Flow cell one was blocked with ethanolamine and used as a control surface. The second one was treated with BSA-Man $\alpha$ 1–3[Man $\alpha$ 1–6] Man (Man-BSA, Dextra) (60  $\mu$ g/ml) in 10 mmol/l acetate buffer, pH 4. Remaining activated groups were blocked with ethanolamine. The final density immobilized on the surface of the second flow cell was 5200 RU. The Man-BSA used to functionalize CM4 chip harbors 12 glycosylation sites. The affinities for DC-SIGN ECD and langerin ECD of pseudo-mannotrioxide and D-mannose were estimated by an inhibition assay, as described [22,31]. Each lectin was injected onto the Man-BSA surface, at 20  $\mu$ mol/l alone or in presence of an increasing concentration of compounds. Injections were performed at 5  $\mu$ l/min using 25 mM Tris-HCl, pH 8, 150 mM NaCl, 4 mmol/l CaCl<sub>2</sub>, and 0.005% of P20 surfactant as running buffer. The surface was regenerated by 1' injection of 50 mmol/l EDTA, pH 8. The IC<sub>50</sub> values were determined as described [22,31].



**Fig. 1. Evaluation of Dendron 12 ability to inhibit HIV-1 BaL *trans* infection in presence of increasing amounts of the virus (a) and selectivity to DC-SIGN (b,c).** (a) After pretreatment with Dendron 12 (250  $\mu$ m) or medium culture, B-THP-1/DC-SIGN cells were pulsed with BaL for 3 h. After washing cells were co-cultured 3 days with CD4<sup>+</sup> T lymphocytes from healthy donors. Levels of infection were quantified by measuring p24 in the supernatants of co-cultures by ELISA. Data were obtained from three different healthy donors. Each donor was tested in duplicate. Values are mean  $\pm$  SD. (b, c) SPR experiment results of DC-SIGN ECD and langerin ECD binding to Man-BSA/dextran surface inhibition by pseudo-mannotriose (psTri) and D-mannose. (b) Inhibition curves, and (c) lectin selectivity histograms.

## Results

### Dendron 12 inhibits *trans* infection in presence of HIV-1 elevated viral load

We firstly examined whether the Dendron 12 was able to inhibit the *trans* infection in the presence of elevated viral load. B-THP1/DC-SIGN cells were used as model to mimic HIV transmission to CD4<sup>+</sup> T cells, as previously described [6,23,32]. B-THP1/DC-SIGN cells were preincubated 30 min in the presence or in absence of Dendron 12 and then exposed to different inoculums of

HIV-1 BaL. Then, B-THP1/DC-SIGN cells were washed and co-cultured with activated CD4<sup>+</sup> T cells. At lower viral concentration inhibition of BaL transmission to CD4<sup>+</sup> T cells was almost complete. Even at higher viral load (40 and 80 TCID<sub>50</sub>) the Dendron 12 was able to counteract the transmission of the virus to CD4<sup>+</sup> T lymphocytes (more than 92% of inhibition) (Fig. 1a).

### Selectivity towards DC-SIGN

Both langerin and DC-SIGN recognize Man<sub>9</sub> on gp120, but have distinct specificities towards complex



oligosaccharides [33–36]. Specificity of pseudo-mannotriose ligand for langerin and DC-SIGN was tested by competition experiments using SPR, as previously described [37]. A CM4 sensor chip was functionalized with BSA-mannotriose and a fixed amount of the extracellular domain of DC-SIGN and langerin was injected over the surface in the presence or absence of pseudo-mannotriose or mannose (control). From the inhibition curves (Fig. 1b), an  $IC_{50}$  of the two ligands towards each lectins was evaluated (Fig. 1c). A limited difference in favor of DC-SIGN was observed for mannose. On the contrary, pseudo-mannotriose is 20 times more potent toward DC-SIGN than against langerin. Moreover, pseudo-mannotriose is 14 times more potent than D-mannose towards DC-SIGN. Indeed, the use of pseudo-mannotriose allows real improvement in affinity and in selectivity.

### Inhibition of HIV-1 infection of human cervical tissue by Dendron 12

Endocervical tissue was obtained from premenopausal women, HIV, HBV and HCV seronegative, undergoing planned therapeutic hysterectomy. Explants were exposed to HIV-1 in a nonpolarized manner, analogous to condition of compromised epithelium *in vivo*. As the laboratory adapted R5 strain HIV-1 BaL is able to infect resting tissue [25], explants were not activated to mimic physiological conditions.

Explants were pretreated 30 min in absence or in presence of increasing concentration of Dendron 12 and then were exposed to BaL in the continued presence of compound. After washing to remove unbound Dendron 12 and virus, explants were maintained in culture up to 7 days. Data represent p24 levels and are presented as the mean of five

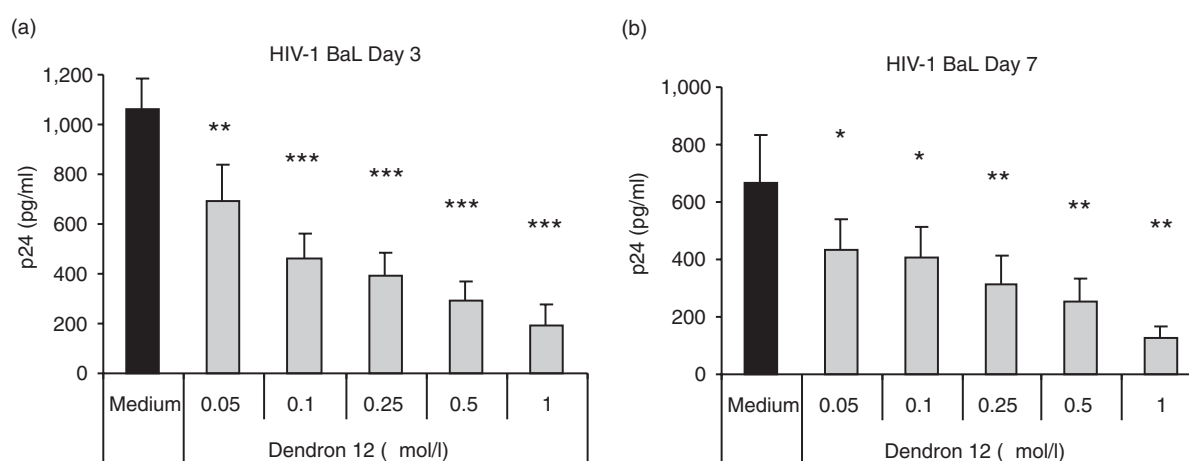
independent experiments, using explants from separate donors. Dendron 12 inhibited cervical explants BaL infection in dose-dependent manner (Fig. 2). At the higher concentration tested, Dendron 12 reduced the infection by about 80%, at 3 and 7 days post infection.

The ability of Dendron 12 to avert explant infection mediated by primary HIV-1 isolates was also verified. R5 tropic isolate 8 g was able to infect unstimulated tissue, but X4 tropic isolate DPMVF needed preactivation to induce infection (not shown). Explants unstimulated or preactivated 2 days with IL-2 and PHA were pretreated with increasing concentration of Dendron 12 and infected, respectively, with 8 g or DPMVF, as described before. Infection inhibition was dose-dependent. At the concentration of 1 mmol/l, infection by both isolates was reduced by more than 85%. At 0.05 mmol/l Dendron 12 decreases by 56% (day 3) and by 40% (day 7) the infection mediated by 8 g (Fig. 3a and b), but the inhibitory effect against DPMVF was largely lost (Fig. 3c and d).

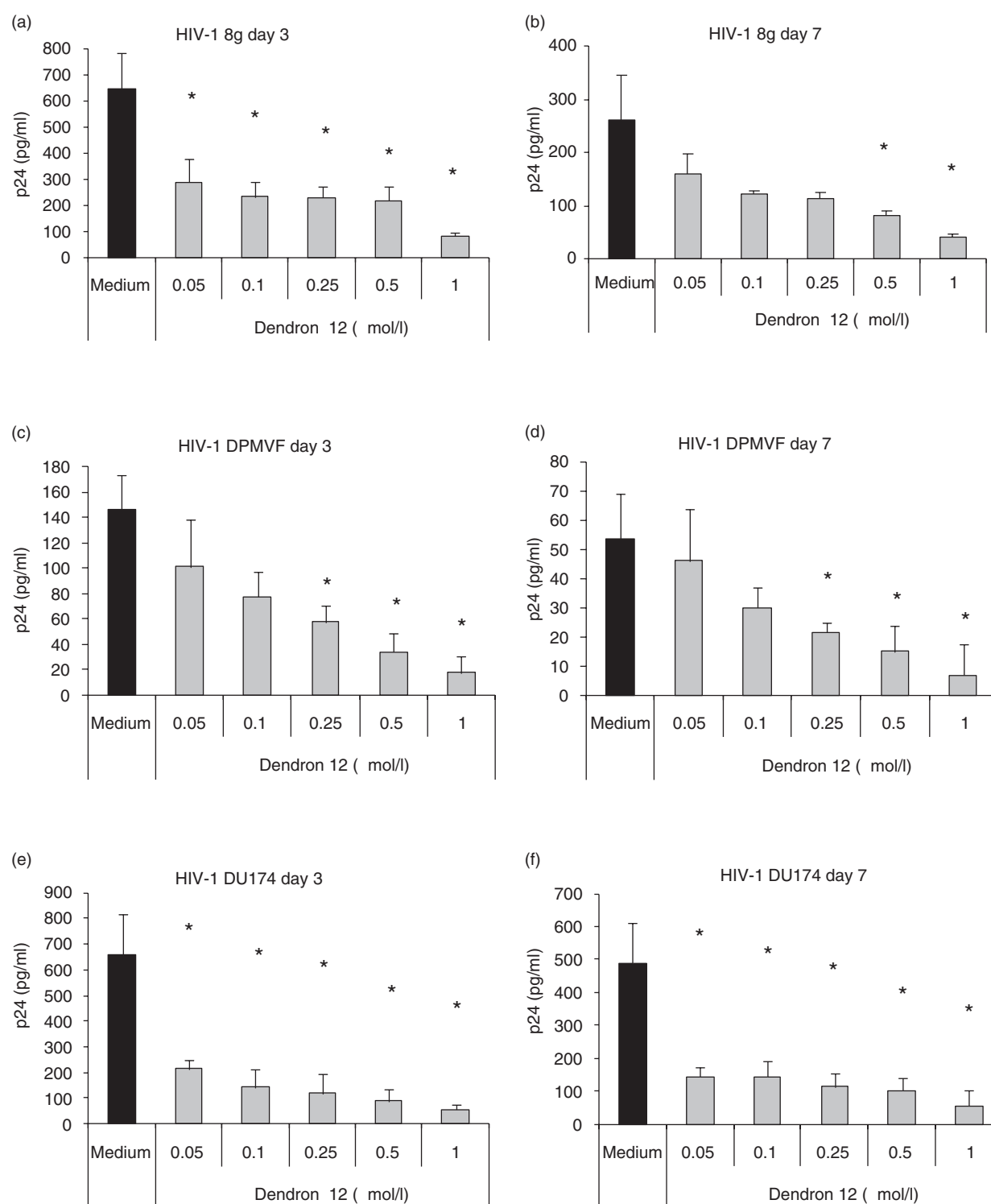
Furthermore we evaluated the capability of Dendron 12 to block explant infection by Clade C R5 tropic strain DU174. Unstimulated explants were pretreated with the compound, exposed to DU174 and cultured as described. Dendron 12 reduced DU174 infection in dose-dependent manner and by about 90% at 1 mM (Fig. 3e and f).

### Induction of $\beta 1$ chemokines production by Dendron 12

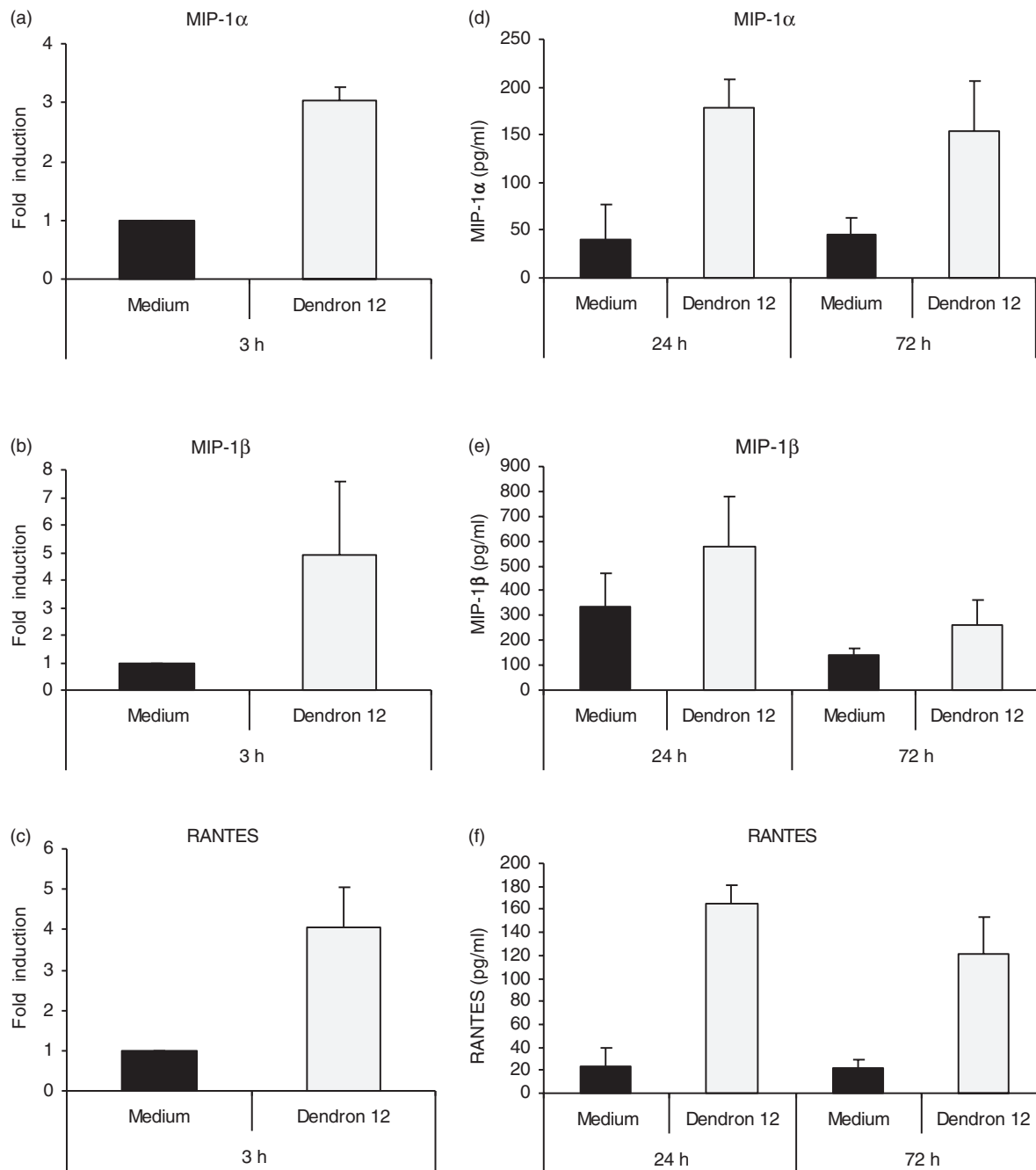
We wondered if, in addition to competitive inhibition of DC-SIGN, other mechanisms account for the antiviral effect of the compound 12. So we investigated if Dendron 12 stimulates the production of factors interfering with



**Fig. 2. Inhibition of endocervical tissue infection mediated by HIV-1 BaL.** After a 30 min pretreatment with the Dendron 12 or medium culture, endocervical explants were incubated 3 h with BaL in the continued presence of indicated concentrations of Dendron 12. Then explants were washed and cultured for 7 days. Infection was monitored by ELISA measurement of p24 in explant culture supernatants at day 3 (a) and 7 (b) post infection. Values represent the mean  $\pm$  SD of five independent experiments. \* $P$  < 0.05 (Student's  $t$ -test), \*\* $P$  < 0.01, \*\*\* $P$  < 0.001.



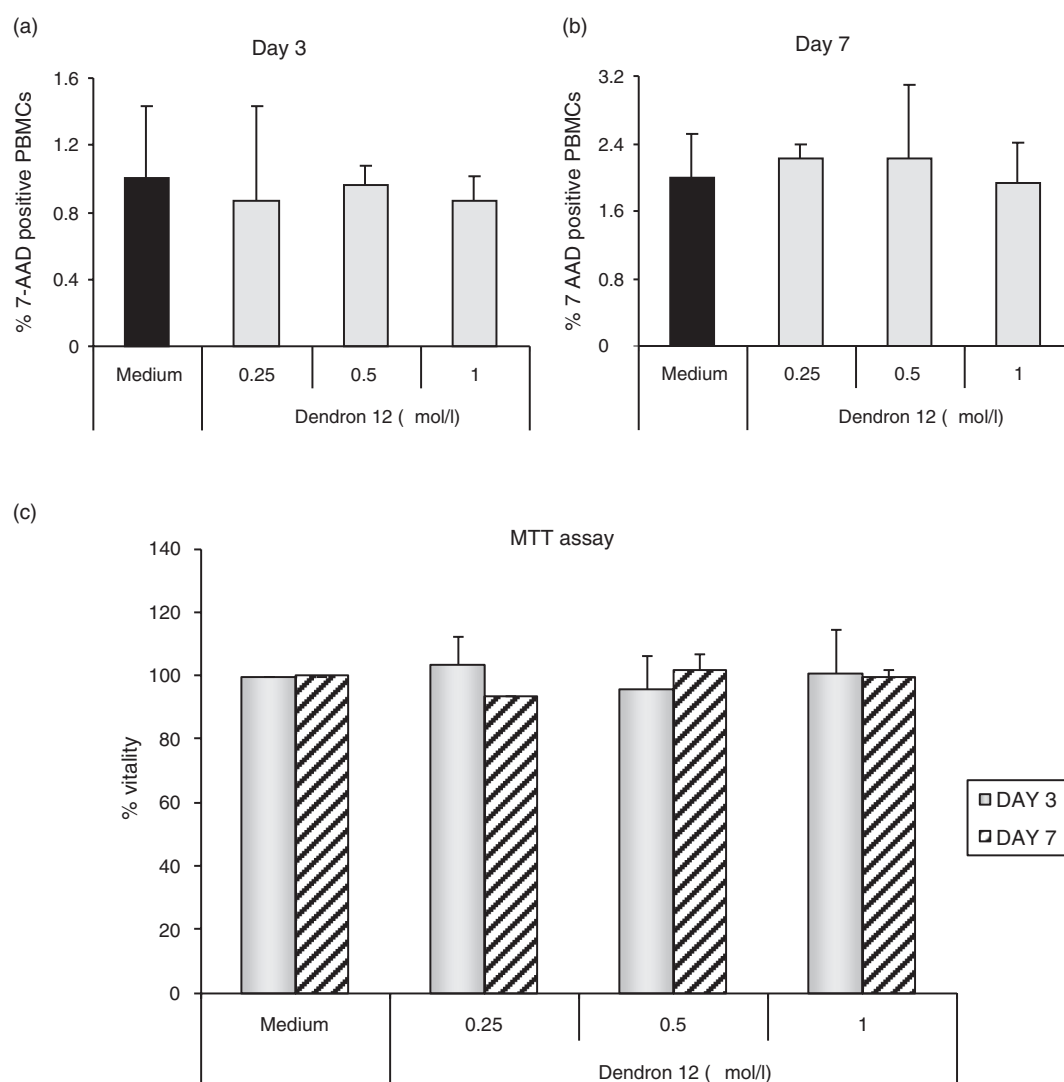
**Fig. 3. Inhibition of endocervical infection induced by HIV-1 primary isolates (8g and DPMVF) and HIV-1clade C DU174.** (a, b, e, f) Unstimulated explants were pretreated 30' with Dendron 12 or medium culture and challenged with 8g or DU174. (c, d) Explants immuno-stimulated were pretreated 30' with Dendron 12 or medium culture and infected with DPMVF. After washing explants were put in culture. Infection was monitored by ELISA measurement of p24 in explant culture supernatants at day 3 and 7 post infection. Values represent the mean  $\pm$  SD of three independent experiments. \* $P < 0.05$  (Student's  $t$ -test).



**Fig. 4.  $\beta$  chemokines induction upon Dendron 12 (250  $\mu$ mol/l) stimulation of human iMDDCs from healthy donors.** (a–c)  $\beta$  chemokines expression following 3 h Dendron 12 (250  $\mu$ mol/l) treatment. mRNA levels of MIP-1 $\alpha$  (a), MIP-1 $\beta$  (b) and RANTES (c) were assessed by quantitative real-time PCR. Expression is normalized to GAPDH and shown as fold changes expression from the unstimulated sample, set as 1. (d–f)  $\beta$  chemokines protein production after stimulation with Dendron 12 (250  $\mu$ mol/l). The concentration of MIP-1 $\alpha$  (d), MIP-1 $\beta$  (e) and RANTES (f) in the culture supernatants was assayed by ELISA at 24 and 72 h. (a–f) Values represent the mean  $\pm$  SD.

HIV infection. Due to difficulty of isolating sufficient amounts of primary mucosal dendritic cells, immature monocyte-derived dendritic cells (iMDDCs) that share with mucosal dendritic cells similar features and DC-SIGN expression, were used as a model [38]. iMDDCs

were treated with Dendron 12 for 3, 24 and 72 h. Expression and production of  $\beta$  chemokines MIP-1 $\alpha$ , MIP-1 $\beta$ , and RANTES following stimulation were analyzed by quantitative real-time PCR and ELISA. The treatment increased expression level of mRNA specific



**Fig. 5. Toxicity of compound 12.** (a, b) Percentage of 7-AAD-positive (nonviable) PBMCs after 3 or 7 days of incubation with different concentrations of Dendron 12 (0.25, 0.5 and 0.1 mmol/l), or in the absence of the inhibitor (medium). Percentage of 7-AAD was determined by flow cytometry. Experiments were performed on PBMCs from three healthy donors. Values are mean  $\pm$  SD. Effect of Dendron 12 on cervical explants viability (c). Endocervical explants, derived from the same donors, were exposed in nonpolarized manner to Dendron 12 or culture medium (control) for 3 and 7 days. Effect of Dendron 12 on tissue viability was determined by the MTT assay. Explants were weighted to normalize optical density of formazan yielded. Viability was expressed as percentage. Experiments were performed on explants from three donors. Data are reported as the average percentage viability ( $\pm$ SD).

for MIP-1 $\alpha$ , MIP-1 $\beta$  and RANTES (Fig. 4a-c) after 3 h. Also MIP-1 $\alpha$ , MIP-1 $\beta$  and RANTES production increased after 24 and 72 h of stimulation (Fig. 4d-f).

#### Evaluation of Dendron 12 toxicity

Cytotoxicity against PBMCs was evaluated by labeling with 7-amino-actinomycin D (7-AAD) that identifies nonviable cells after Dendron 12 treatment. Compound exposure for 3 and 7 days did not alter significantly the viability of PBMCs (Fig. 5a and b).

To assess toxicity of Dendron 12 towards cervical tissue, effect of nonpolarized exposure of the compound on explants viability was monitored (Fig. 5c). After 3 and 7 days treatment in absence or in presence of different concentration of Dendron 12, viability was evaluated by a MTT-based assay. Viability of the compound treated explants was compared to viability of untreated control. No significant difference between control and treated explants was observed up to a concentration of 1 mM (the higher concentration tested in assessment of compound efficacy against HIV-1 infection).

## Discussion

Three decades after HIV discovery HIV-AIDS pandemic continues and millions of people are infected every year. Thus, the development of effective, nontoxic and low-cost topical microbicides represents a valid alternative approach to prevent the sexual transmission of HIV [3]. However, so far almost all compounds failed to prevent HIV transmission in efficacy trials. A recent exception was a vaginal gel formulation of tenofovir that reduced HIV infection by 50% [39].

Tenofovir and other topical microbicides can prevent localized infection of target cells in genital mucosae. However, dendritic cells DC-SIGN<sup>+</sup> internalize and transport HIV to secondary lymphoid organs, rendering the virus inaccessible to inhibitory effect of the microbicides. Therefore inhibition of DC-SIGN is essential to block HIV-1 uptake and dissemination from migratory dendritic cells.

We have previously reported that the tetravalent Dendron **12** was able to block almost completely the HIV-1 *trans* infection of CD4<sup>+</sup> T cells at micromolar range [23]. This compound exerts its activity by competitive inhibition of HIV-1 gp120 binding to DC-SIGN. In our initial experiments we demonstrated that Dendron **12** (at the same concentration previously assayed) even in presence of higher viral loads retains its ability to inhibit HIV-1 *trans* infection. This potent inhibitory activity is due both to elevate affinity for DC-SIGN of Dendron **12** pseudo-trisaccharide units and to high avidity of binding, guaranteed by the tetravalent presentation on the compound scaffold.

Different DC-SIGN inhibitors have been described so far. Dendrons displaying complex oligomannoses in high density inhibited binding of gp120 to DC-SIGN with IC<sub>50</sub> in nanomolar range [40]. Nevertheless the complexity of the oligosaccharides used limits the possibility of a therapeutic application. Gold nanoparticles displaying mannosyl oligosaccharides are potent inhibitors of DC-SIGN-mediated HIV-1 *trans* infection of human PBMCs [41], but may have toxic effects because of gold accumulation.

Unlike DC-SIGN, the C-type lectin langerin, expressed on Langerhans cells, appears to play a protective role against HIV infection. Langerhans cells are located in the stratified mucosal epithelia of female and male genital tissue. Both DC-SIGN and langerin recognize high mannose glycans on gp120 through their carbohydrate recognition domain (CRD). Recent data demonstrate that langerin prevents HIV transmission by Langerhans cells, at least in the presence of low concentration of the virus, promoting rapid degradation and clearance of HIV-1 [32,33]. Using a biosensor with a SPR detection method we showed that the Dendron **12** is much more

selective for DC-SIGN than for langerin. Considering the similarity between the CRD of DC-SIGN and langerin, the fact that the Dendron **12** does not interfere with langerin function is a remarkable advantage in view of developing new microbicides.

The efficacy and the safety of the Dendron **12** were evaluated in a human cervical explant model. Endocervix is more susceptible to HIV-1, being lined by a single layer of columnar epithelium, and contains dendritic cell DC-SIGN<sup>+</sup> in the subepithelial region [4,5,24]. Explants were exposed to HIV-1 to mimic a condition of compromised epithelium *in vivo*; condition that highly increases the risk of HIV infection.

R5-tropic virus strains predominate during HIV-1 transmission *in vivo* and cause the majority of new infections [42]. The results obtained showed that Dendron **12** strongly reduced the infection of cervical explants by different HIV-1 R5 tropic strains, such as BaL, the primary isolate 8 g and the Clade C DU174. This may have a great impact, considering that Clade C is the most abundant subtype in all countries of Southern Africa and in some countries of eastern Africa and Asia, areas where the majority of HIV-1-infected people resides [43,44].

Rare cases of HIV-1 infection by X4-tropic strains were observed in CCR5Δ32 homozygotes and X4-tropic strain can infect immune activate human cervical tissue [25,42]. Dendron **12** was able to inhibit DPMVF primary X4-tropic strain infection at higher concentration assayed, but this effect was in part lost at lower concentrations of the compound.

Data obtained suggest that Dendron **12** might have additional mechanisms of action apart from competitive inhibition of HIV-1 binding to DC-SIGN. Flow cytometric studies indicate that treatment with Dendron **12** reduces DC-SIGN expression on B-THP1/DC-SIGN<sup>+</sup> cells [23]. This effect may be due to increased internalization of the receptor after binding of the compound to DC-SIGN.

Furthermore, Dendron **12** induced an increase of the production of β chemokines, such as MIP-1α, MIP-1β and RANTES, by iMDDCs. These β chemokines, natural ligands of CCR5, suppress HIV-1 R5 tropic strain replication, competing with the binding of the virus to CCR5 or inducing receptor internalization [45–47]. The Dendron **12**, enhancing β chemokines production, could interfere, at least partially, with the direct infection of CCR5<sup>+</sup> CD4<sup>+</sup> T lymphocytes and macrophages located in genital mucosae.

The Dendron **12** does not reduce the vitality of PBMCs and of explants. However, additional experiments, such as rabbit vaginal irritation assay [48], would be needed for a more accurate evaluation of potential toxic effects.

Vaginal epithelium has limited permeability to particles greater than 30 nm [2]. However, the diameter of Dendron **12** is certainly below that threshold, so the compound could enter and diffuse into intact mucosal tissue. Furthermore the Dendron **12** scaffold can be easily modified to improve absorption of the compound, without decreasing affinity to DC-SIGN.

The tetravalent Dendron **12** prevents HIV *trans* infection of CD4<sup>+</sup> T lymphocytes at micromolar range, even in presence of elevated viral load, and displays high solubility in physiological media, a neglectable toxicity and a long-lasting effect. Moreover it inhibits in dose-dependent manner HIV-1 infection of human cervical explants. These features make the Dendron **12** a good candidate as a lead compound to develop new microbicide drugs.

However, the Dendron **12** inhibits 80–90% of cervical explants HIV-1 infection, but cannot block it completely. To overcome these limitations, the structure of this compound can be improved in both the scaffold and the active pseudo-saccharide ligand to develop new ligands of DC-SIGN more effective and easier to synthesize. Furthermore, Dendron **12** (or its derivatives) can be used in combination with other molecules directed against different HIV targets. In particular multivalent structures, presenting on the same scaffold multiple copies of DC-SIGN and HIV co-receptor inhibitors, simultaneously blocking different HIV targets, could protect against different routes of HIV transmission.

## Acknowledgements

M.C., A. Bernardi, D.T. conceived the study; A. Berzi, M.C. wrote the paper, A. Berzi, M.B. performed the experiments and analyzed the data; Jo.Re., R.O. synthesized the pseudo-mannotriose; Ja.Ro., M.S.N. synthesized the Dendron scaffold; F.F., I.S. performed SPR experiments and analyzed the data; E.C. expressed and purified DC-SIGN and langerin; I.C., P.A. obtained cervical explants.

## Conflicts of interest

Conflict of interest and source of funding: the authors declare no conflict of interest.

B-THP-1 and B-THP-1/DC-SIGN cells, HIV-1 BaL and DU174 were provided through the EU programme EVA centre for AIDS Reagents NIBSC, UK. Buffy coats from healthy donors were provided by the Transfusional Unit of Vimercate Hospital (Italy). This work was supported by: the FIRB program CEM-PROFAR-MANET (RBPR05NWWC), Istituto Superiore di Sanità 'Programma Nazionale di Ricerca sull'AIDS', EMPRO and AVIP EC WP6 Projects, the nGIN EC WP7 Project, 2008 Ricerca Finalizzata and Ricerca

Corrente (Italian Ministry of Health), the MICINN of Spain CTQ2008-01694/BQU European FP7 project: EU ITN CARMUSYS (PITN-GA-2008-213592), Sidaction – ensemble contre le sida. A. Berzi was supported by a fellowship of Doctorate School of Molecular Medicine, University of Milan; I.S. by ITN Carmusys, E.C. by a fellowship from 'ministère de la recherche et de l'enseignement supérieur'; M.S.N. by FPU fellowship.

## References

- UNAIDS Report on the global AIDS epidemic, 2010.
- Cutler B, Justman J. **Vaginal microbicides and the prevention of HIV infection.** *Lancet Infect Dis* 2008; **8**:685–697.
- Reina JJ, Bernardi A, Clerici M, Rojo J. **HIV microbicides: state-of-the-art and new perspectives on the development of entry inhibitors.** *Future Med Chem* 2010; **2**:1141–1159.
- Geijtenbeek TBH, Torensma R, van Vliet SJ, van Duijnhoven GCF, Adema GJ, van Kooyk Y, et al. **Identification of DC-SIGN, a novel dendritic cell-specific ICAM-3 receptor that supports primary immune responses.** *Cell* 2000; **100**:575–585.
- Pope M, Haase AT. **Transmission, cute HIV infection and the quest for strategies to prevent infection.** *Nat Med* 2003; **9**:847–852.
- Wu L, Wu L, KevalRamani VL. **Dendritic-cell interactions with HIV: infection and viral dissemination.** *Nat Rev Immunol* 2006; **6**:859–868.
- Geijtenbeek TBH, Kwon DS, Torensma R, van Vliet SJ, van Duijnhoven GCF, Middel J, et al. **DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T Cells.** *Cell* 2000; **100**:587–597.
- Van Kooyk Y, Geijtenbeek TBH. **DC-SIGN: escape mechanism for pathogens.** *Nat Rev Immunol* 2003; **3**:697–709.
- Kwon DS, Gregorio G, Bitton N, Hendrikson WA, Littman DR. **DC-SIGN-mediated internalization of HIV is required for trans-enhancement of T cell infection.** *Immunity* 2002; **16**:135–144.
- McDonald D, Wu L, Bohks SM, KewalRamani VN, Unutmaz D, Hope TJ. **Recruitment of HIV and its receptors to dendritic cell–T cell junctions.** *Science* 2003; **300**:1295–1297.
- Wiley RD, Gummuluru S. **Immature dendritic cells derived exosomes can mediate HIV-1 trans infection.** *Proc Natl Acad Sci USA* 2006; **103**:738–743.
- Lee B, Leslie G, Soilleux E, O'Doherty U, Baik S, Levroney E, et al. **Cis expression of DC-SIGN allows for more efficient entry of human and simian immunodeficiency viruses via CD4 and a co-receptor.** *J Virol* 2001; **75**:12028–12038.
- Turville SG, Santos JJ, Frank I, Cameron PU, Wilkinson J, Miranda-Saksena M, et al. **Immunodeficiency virus uptake, turnover, and 2-phase transfer in human dendritic cells.** *Blood* 2003; **103**:2170–2179.
- Gringhuis SI, den Dunnen J, Litjens M, van Het Hof B, van Kooyk Y, Geijtenbeek TBH. **C-type lectin DC-SIGN modules Toll-like receptor signaling via Raf-1 kinase-dependent acetylation of transcription factor NF-κB.** *Immunity* 2007; **26**:605–616.
- Hodges A, Sharrocks K, Edelmann M, Baban D, Moris A, Schwartz O, et al. **Activation of the lectin DC-SIGN induces an immature dendritic cell phenotype triggering Rho-GTPase activity required for HIV-1 replication.** *Nat Immunol* 2007; **8**:569–577.
- Gringhuis SI, van der Vlist M, van den Berg LM, den Dunnen J, Litjens M, Geijtenbeek TBH. **HIV-1 exploits innate signaling by TLR8 and DC-SIGN for productive infection of dendritic cells.** *Nat Immunol* 2010; **11**:419–426.
- Gringhuis SI, den Dunnen J, Litjens M, van der Vlist M, Geijtenbeek TBH. **Carbohydrate-specific signaling through the DC-SIGN signalosome tailors immunity to Mycobacterium tuberculosis, HIV-1 and Helicobacter pylori.** *Nat Immunol* 2010; **10**:1081–1089.

18. Feinberg H, Mitchell DA, Drickamer K, Weis WI. **Structural basis for selective recognition of oligosaccharides by DC-SIGN and DC-SIGNR.** *Science* 2001; **294**:2163–2166.
19. Adams EW, Ratner DM, Bokesch HR, McMahon JB, O’Keefe BR, Seeburger PH. **Oligosaccharide and glycoprotein microarrays as tools in HIV glycobiology; glycan-dependent gp120/protein interactions.** *Chem Biol* 2004; **11**:875–881.
20. Reina JJ, Sattin S, Invernizzi D, Mari S, Martínez-Prats L, Tabarani G, *et al.* **1,2 Mannobioside mimic: synthesis, DC-SIGN interaction by NMR and docking, and antiviral activity.** *Chem MedChem* 2007; **2**:1030–1036.
21. Rojo J, Delgado R. **Glycodendritic structures: promising new antiviral drugs.** *J Antimicrob Chemother* 2004; **54**:579–581.
22. Luczkowiak J, Sattin S, Sutkevičiūtė I, Reina JJ, Sánchez-Navarro M, Thépaut M, *et al.* **Pseudosaccharide functionalized dendrimers as potent inhibitors of DC-SIGN dependent Ebola pseudotyped viral infection.** *Bioconj Chem* 2011; **22**:1354–1365.
23. Sattin S, Daggetti A, Thépaut M, Berzi A, Sanchez-Navarro M, Rojo J, *et al.* **Inhibition of DC-SIGN-mediated infection by a linear trimannoside mimic in a tetravalent presentation.** *ACS Chem Biol* 2010; **5**:301–312.
24. Anderson DJ, Pudney J, Shust DJ. **Caveat associated with the use of human cervical tissue for HIV and microbicide research.** *AIDS* 2010; **24**:1–4.
25. Greenhead P, Hayes P, Watts P, Laing K, Griffin G, Shattock R. **Parameters of Human Immunodeficiency Virus Infection of human cervical tissue and inhibition by vaginal virucides.** *J Virol* 2000; **74**:5577–5586.
26. Wallace GS, Cheng-Mayer C, Schito ML, Fletcher P, Miller Jenkins LM, Hayashi R, *et al.* **Human immunodeficiency virus type 1 nucleocapsid inhibitors impede trans infection in cellular and explant models and protect nonhuman primates from infection.** *J Virol* 2009; **83**:9175–9182.
27. Cummins JE, Guarner J, Flowers L, Guenther PC, Bartlett J, Morken T, *et al.* **Preclinical testing of candidate topical microbicides for antihuman immunodeficiency virus type 1 activity and tissue toxicity in a human cervical explant culture.** *Antimicrob Agents Chemother* 2007; **51**:1770–1779.
28. Thépaut M, Valladeau J, Nurisso A, Kahn R, Arnou B, Vivès C, *et al.* **Structural studies of Langerin and Birbeck granule: a macromolecular organization model.** *Biochemistry* 2009; **48**:2684–2698.
29. Tabarani G, Thépaut M, Stroebel D, Ebel C, Vivès C, Vachette P, *et al.* **DC-SIGN neck domain is a pH-sensor controlling oligomerization: SAXS and hydrodynamic studies of extracellular domain.** *J Biol Chem* 2009; **284**:21229–21240.
30. Halary F, Amara A, Lortat-Jacob H, Messerle M, Delaunay T, Houlès C, *et al.* **Human cytomegalovirus binding to DC-SIGN is required for dendritic cell infection and target cell trans-infection.** *Immunity* 2002; **17**:653–664.
31. Andreini M, Doknic D, Sutkevičiūtė I, Reina JJ, Duan J, Chabrol E, *et al.* **Second generation of fucose-based DC-SIGN ligands: affinity improvement and specificity versus Langerin.** *Org Biomol Chem* 2011; **9**:5778–5786.
32. Wu L, Martin TD, Carrington M, KewalRamani VN. **Raji B cells, misidentified as THP-1 cells, stimulate DC-SIGN-mediated HIV transmission.** *Virology* 2004; **318**:17–23.
33. de Witte L, Nabatov A, Pion M, Fluittsma D, Marein AW, de Jong P, *et al.* **Langerin is natural barrier to HIV-1 transmission by langerhans cells.** *Nat Med* 2007; **13**:367–371.
34. de Witte L, Nabatov A, Geijtenbeek TBH. **Distinct roles for DC-SIGN-dendritic cells and langerhans cells in HIV-1 transmission.** *Trends Mol Med* 2007; **14**:12–19.
35. Guo Y, Feinberg H, Conroy E, Mitchell DA, Alvarez R, Blixt O, *et al.* **Structural basis for distinct ligand-binding and targeting properties of the receptors DC-SIGN and DC-SIGNR.** *Nat Struct Mol Biol* 2004; **11**:591–598.
36. Galustian C, Park CG, Chai W, Kiso M, Bruening SA, Kang YS, *et al.* **High and low affinity carbohydrate ligands revealed for murine SIGN-R1 by carbohydrate array and cell binding approaches, and differing specificities for SIGN-R3 and Langerin.** *Int Immunol* 2004; **16**:853–866.
37. Timpano G, Tabarani G, Anderlueh M, Invernizzi D, Vasile F, Potenza D, *et al.* **Synthesis of novel DC-SIGN ligands with an alpha-fucosylamide anchor.** *ChemBioChem* 2008; **9**:1921–1930.
38. Sallusto F, Lanzavecchia A. **Efficient presentation of soluble antigen by cultured human dendritic cells is maintained by granulocyte/macrophage colony-stimulating factor plus interleukin 4 and downregulated by tumor necrosis factor alpha.** *J Exp Med* 1994; **179**:1109–1118.
39. Abdool Karim Q, Abdool Karim SS, Frohlich JA, Grobler AC, Baxter C, Mansoor LE, *et al.* **Effectiveness and safety of tenofovir gel, an antiretroviral microbicide, for the prevention of HIV infection in women.** *Science* 2010; **329**:1168–1174.
40. Wang SK, Liang PH, Astronomo RD, Hsu TL, Hsieh SL, Burton DR, *et al.* **Targeting the carbohydrates on HIV-1: interaction of oligomannose dendrons with human monoclonal antibody 2G12 and DC-SIGN.** *Proc Natl Acad Sci USA* 2008; **105**:3690–3695.
41. Martinez-Avila O, Bedoya LM, Marradi M, Clavel C, Alcami J, Penades S. **Multivalent manno-glycanoparticles inhibit DC-SIGN-mediated HIV-1 trans-infection of human T cells.** *Chem Biochem* 2009; **10**:1806–1809.
42. Nazari R, Joshi S. **CCR5 as target for HIV-1 gene therapy.** *Curr Gene Ther* 2008; **8**:1–9.
43. McCutchan FE. **Understanding the genetic diversity of HIV-1.** *AIDS* 2000; **14** (Suppl 3):S31–S44.
44. Spira S, Wainberg MA, Loemba H, Turner D, Brenner BG. **Impact of clade diversity on HIV-1 virulence, antiretroviral drug sensitivity and drug resistance.** *J Antimicrob Chemother* 2003; **51**:229–240.
45. Samson M, Libert F, Doranz BJ, Rucker J, Liesnard C, Farber CM, *et al.* **Resistance to HIV-1 infection in Caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene.** *Nature* 1996; **26**:722–725.
46. Wu L, LaRosa G, Kassam N, Gordon CJ, Heath H, Ruffing N, *et al.* **Interaction of chemokine receptor CCR5 with its ligands: multiple domains for HIV-1 gp120 binding and a single domain for chemokine binding.** *J Exp Med* 1997; **186**:1373–1381.
47. Cocchi F, DeVico AL, Garzino-Demo A, Arya SK, Gallo RC, Lusso P. **Identification of RANTES, MIP-1 $\alpha$ , and MIP-1 $\beta$  as the major HIV-1-suppressive factors produced by CD8<sup>+</sup> T cells.** *Science* 1995; **270**:1811–1815.
48. Eckstein P, Jackson MC, Millman N, Sobrero AJ. **Comparison of vaginal tolerance tests of spermicidal preparations in rabbits and monkeys.** *J Reprod Fertil* 1969; **20**:85–93.



## Second generation of fucose-based DC-SIGN ligands : affinity improvement and specificity *versus* Langerin†

Manuel Andreini,<sup>‡a</sup> Daniela Doknic,<sup>‡a</sup> Ieva Sutkeviciute,<sup>‡b,e</sup> José J. Reina,<sup>a</sup> Janxin Duan,<sup>c</sup> Eric Chabrol,<sup>b,e</sup> Michel Thepaut,<sup>b,e,f</sup> Elisabetta Moroni,<sup>a</sup> Fabio Doro,<sup>a</sup> Laura Belvisi,<sup>a</sup> Joerg Weiser,<sup>c</sup> Javier Rojo,<sup>d</sup> Franck Fieschi<sup>\*b,e,g</sup> and Anna Bernardi<sup>\*a</sup>

Received 11th April 2011, Accepted 20th May 2011

DOI: 10.1039/c1ob05573a

DC-SIGN and Langerin are two C-type lectins involved in the initial steps of HIV infections: the former acts as a viral attachment factor and facilitates viral invasion of the immune system, the latter has a protective effect. Potential antiviral compounds targeted against DC-SIGN were synthesized using a common fucosylamide anchor. Their DC-SIGN affinity was tested by SPR and found to be similar to that of the natural ligand Lewis-X (Le<sup>x</sup>). The compounds were also found to be selective for DC-SIGN and to interact only weakly with Langerin. These molecules are potentially useful therapeutic tools against sexually transmitted HIV infection.

### Introduction

Dendritic Cells (DCs) are instrumental in the development of pathogen-specific immune responses.<sup>1</sup> DCs are professional antigen-presenting cells that capture microbes entering skin or mucosal tissues and process them to form MHC-peptide complexes. After antigen uptake, immature DCs acquire the capacity to migrate to lymph nodes where they present processed antigens to T-cells, initiating adaptive immune responses. DCs express a repertoire of pathogen-recognition receptors (PRRs), including Toll Like Receptors (TLRs) and C-type lectins that mediate both signaling by self antigens and, in some cases, pathogen recognition.<sup>2</sup> C-type lectins represent a large family of Ca<sup>2+</sup> dependent lectins and recognize pathogen-derived carbohydrate structures. Many different C-type lectins expressed by DCs have been described,<sup>3</sup> including DC-SIGN.

DC-SIGN (DC-Specific ICAM-3 Grabbing Nonintegrin; CD 209) was originally defined as an intercellular adhesion molecule-3

(ICAM-3) receptor that plays an important role in establishing the first contact between DC and resting T cells.<sup>4</sup> It is a type II trans-membrane C-type lectin with a single C-terminal Carbohydrate Recognition Domain (CRD) within its sequence. In the cellular membrane, DC-SIGN is assembled as a tetramer, thanks to an extended coiled-coil region that allows simultaneous presentation of four CRDs.<sup>5</sup> This oligomerization influences the lectin avidity in binding events. DC-SIGN appears to promote dissemination of a number of viruses (e.g., HIV, hepatitis C virus, Ebola virus)<sup>6</sup> and to participate in suppressing immune responses to some pathogens, (e.g., *Mycobacterium tuberculosis* and *Helicobacter pylori*).<sup>7</sup>

The various roles attributed to DC-SIGN have generated much interest towards the identification of ligands that can be used to explore its different functions and/or to inhibit pathogen binding.<sup>8</sup> However, generation of specific ligands for DC-SIGN is a challenging task, since many other C-type lectins exist and share important structural features with DC-SIGN binding site. Among the list of C-type lectin receptors closely related to DC-SIGN, Langerin, which is also expressed at the cell surface of antigen presenting cells, and L-SIGN, expressed on endothelial liver cell, placenta and lymph nodes, are particularly likely to interfere with DC-SIGN recognition.<sup>9,10</sup> These three lectins are all calcium-dependent carbohydrate-binding proteins and share the ability to bind high-mannose oligosaccharides. On the other hand, the three lectins show different specificity towards fucosylated oligosaccharides, a fact which may be used to design DC-SIGN specific ligands. Indeed, contrary to L-SIGN, DC-SIGN is known to bind the Lewis X (Le<sup>x</sup>) epitope (Galβ4[Fucα3]GlcNAc, **1** in Fig. 1), as illustrated in the recognition mode of *Schistosoma mansoni* egg by these lectins.<sup>5d,7,11</sup> Moreover, both DC-SIGN and Langerin appear to recognize blood group B antigen through its fucose residue<sup>12,13</sup> in the primary Ca<sup>2+</sup> binding site, but again the Le<sup>x</sup> antigen is specific for DC-SIGN relative to Langerin. As

<sup>a</sup>Università degli Studi di Milano, Dipartimento di Chimica Organica e Industriale and CISI, via Venezian 21, 20133 Milano, Italy. E-mail: anna.bernardi@unimi.it

<sup>b</sup>Institut de Biologie Structurale, Université Grenoble I, 41 rue Jules Horowitz, 38027 Grenoble, France. E-mail: fieschi@ibs.fr

<sup>c</sup>Anterio consult&research, Augustaanlage 23, D-68165 Mannheim

<sup>d</sup>Glycosystems Laboratory, Instituto de Investigaciones Químicas, CSIC–Universidad de Sevilla, Américo Vespucio 49, 41092 Sevilla, Spain

<sup>e</sup>CNRS, UMR 5075, Grenoble, France

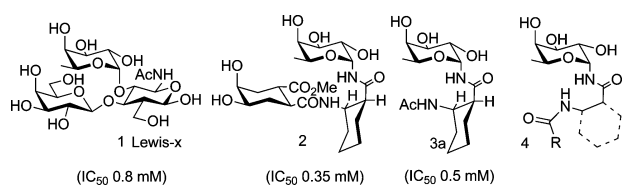
<sup>f</sup>CEA, Grenoble, France

<sup>g</sup>Institut Universitaire de France, 103 boulevard Saint-Michel, 75005 Paris, France

† Electronic supplementary information (ESI) available: . See DOI: 10.1039/c1ob05573a

‡ These authors, listed in alphabetical order, contributed equally to this work





**Fig. 1** The known fucose-based DC-SIGN ligands **1–3a** and the general structure **4** of the library described in this paper. ( $IC_{50}$  from ref. 14).

shown by X-ray structures,<sup>12,13</sup> in addition to binding the fucose residue on the  $Ca^{2+}$  site, DC-SIGN is uniquely able to stabilize  $Le^x$  galactose residue in a second binding area due to key residues that are absent in Langerin.

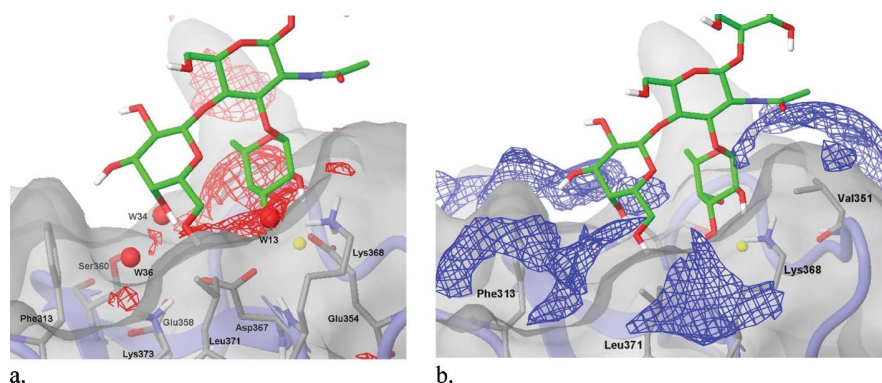
The existence of a secondary binding site is also suggested by a glycan array study of over 100 glycan structures.<sup>12</sup> This study demonstrated that the presence of a terminal fucose residue is not sufficient for DC-SIGN binding, but 14 fucose-bearing glycans with the structure of Lewis epitopes were found to bind selectively to DC-SIGN relative to L-SIGN.

We have recently described the first fucose-based artificial ligand of DC-SIGN (compound **2**, Fig. 1), designed to mimic this trisaccharide.<sup>14</sup> The ligand was built by using an  $\alpha$ -fucosylamide anchor to drive the molecule to the DC-SIGN primary binding site and connecting it to a galactose mimic using a cyclic *cis*- $\beta$ -amino acid ((1*S*,2*R*)-2-amino-cyclohexanecarboxylic acid, Fig. 1). The layout of these residues and more specifically the linker  $\beta$ -amino acid allowed the molecule to adopt a three-dimensional shape similar to the  $Le^x$  trisaccharide (Fig. 1).<sup>15</sup> Amide bonds were chosen to connect the three elements of ligand **2** to achieve synthetic simplicity as well as chemical and metabolic stability of the target molecule. DC-SIGN binding studies performed by SPR showed that ligand **2** and surprisingly its simplified version **3a**, which does not contain the galactose-mimic moiety, inhibit DC-SIGN better than the natural ligand **1** ( $IC_{50}$  = 0.35 mM, 0.5 mM and 0.8 mM, respectively).<sup>14</sup> The weak difference of affinity between **2** and **3a**, however, suggested that the galactose-like fragment in **2** gives a limited contribution to the binding interaction. Building on this knowledge, the goals of the work we report in this paper were: 1) to establish a minimal structure easily accessible in large scale and able to engage the receptor with an affinity similar to that of the natural ligand  $Le^x$ ; 2) to

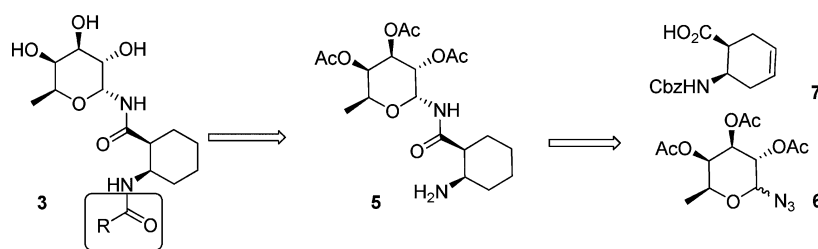
improve the binding affinity of the fucosylamides by optimizing the interactions in the secondary binding site. To achieve these goals, a library of *ca.* 40 derivatives of general formula **4** was designed, synthesized and assayed by SPR to determine the ability of the compounds to inhibit DC-SIGN binding to immobilized mannosylated Bovine Serum Albumin (Man-BSA). Moreover, a preliminary selectivity screening was introduced to test some library members for inhibition of Langerin, using SPR. Selectivity for DC-SIGN *versus* Langerin is specially important to develop inhibitors of sexually transmitted HIV infections. As discussed above, interaction with DC-SIGN on mucosal DC is used by the virus to invade the host immune system. On the contrary, Langerin is suggested to have protective effects against HIV infection.<sup>16</sup> Indeed, some of the fucosylamides examined displayed an interesting DC-SIGN selectivity and have the potential of being developed as antiviral agents.

## Results and discussion

In order to select reasonable ligand candidates, the properties of the protein surface in the vicinity of the  $Ca^{2+}$  site in the Lewis<sup>x</sup>-DC-SIGN complex (1SL5)<sup>12</sup> were examined using GRID.<sup>17</sup> Both the DRY probe and the WATER probe were used to identify hydrophilic and lipophilic regions of the binding site. The molecular representations shown in Fig. 2 were obtained with the Maestro graphical interface. Various minima for the WATER probe were identified in the vicinity of the  $Ca^{2+}$ -binding region: in the crystal structure, they are occupied by crystallographic water molecules W13, W34 and W36 (Fig. 2a). W13 and W36 are located in two well-defined low interaction energy sites, both below  $-11$  kcal mol<sup>-1</sup>. W13 is in the vicinity of the fucose residue and mediates the interaction of Fuc-O2 with Glu354 and Lys368, and of Gal-O6 with Asp367 (Fig. 2a). W36 mediates the interaction of Gal-O4 with Glu358. The W36 site is occupied by a crystallographic water molecule also in 3 out of 4 known X-ray structures of DC-SIGN in complex with oligomannosides (1SL4,<sup>12</sup> 1K9I,<sup>18</sup> 2IT5<sup>19</sup>) and it is replaced by one sugar hydroxyl group in the fourth one (2IT6<sup>19</sup>). The W34 site belongs to a larger isoenergetic area with a less favorable GRID interaction energy (*ca.*  $-7$  kcal mol<sup>-1</sup>), occupied by water molecules in 2 out of 4 oligomannoside-DC-SIGN complexes (1SL4, 1K9I) and loosely



**Fig. 2** Energetic maps of GRID interactions (the  $Ca^{2+}$  ion is shown in yellow) in the structure of the DC-SIGN-  $Le^x$  complex (pdb code : 1SL5, from ref. 11). a) WATER probe,  $-7.0$  kcal mol<sup>-1</sup> isosurface, showing the binding sites for crystallographic water molecules W13, W34 and W36 in 1SL5. b) DRY probe,  $-0.5$  kcal mol<sup>-1</sup> isosurface, showing hydrophobic areas near the binding region and the groove formed by Phe313 and Leu371.

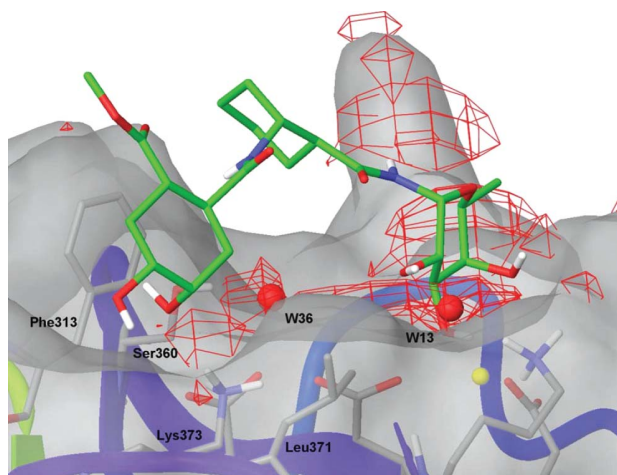


**Scheme 1** Retrosynthetic analysis of fucosyl derivative **3**.

replaced by mannose hydroxyl groups in the other **2** (2IT5 and 2IT6). In the 1SL5 structure, W34 mediates binding of Gal-O4 to Ser360 and, together with W36, contributes to the creation of a secondary binding site involving Leu371, Asp367, Lys373, Glu358 and flanked by Phe313 (Fig. 2a).

GRID analysis with the DRY probe allowed identification of the hydrophobic areas near the binding region, which are shown in Fig. 2b. Two of them, formed by Val351 and Asn362/Asn344, are in the immediate vicinity of the  $\text{Ca}^{2+}$ -binding site and establish Van der Waals contact with the ligand. Phe313 and Leu371/Lys368 side chains form a major hydrophobic groove which flank the W36 crystallographic site.

Docking of mimic **2** in the 1SL5 structure was obtained using Glide.<sup>20</sup> The complex, which included protein, ligand and the two water molecules W13 and W36, was prepared with the standard Preparation Wizard routine of Glide, but the final minimization was performed in implicit (GB/SA<sup>21</sup>) water with the AMBER\* force field. It was found that this procedure allowed achievement of a better orientation of the water molecules and optimization of their hydrogen bonding pattern, which in turn avoided steric clashes in the following re-docking step. Docking obtained with this model suggested that optimal interaction is reached with the ligand in an extended conformation, which would allow the galactose mimic to place two hydroxyl groups in a hydrophilic patch of the protein near the side chain of Phe313 while nesting the cyclohexane ring in the groove formed by Phe313 and Leu371 (Fig. 3).



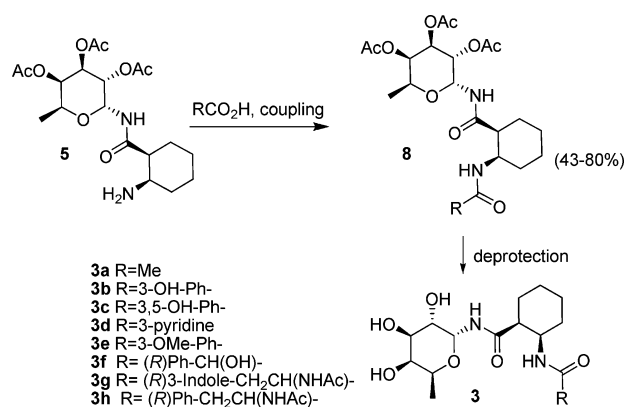
**Fig. 3** Docking of mimic **2** in 1SL5. (the  $\text{Ca}^{2+}$  ion is shown in yellow).

Based on this model, interesting candidates to replace the galactose moiety should be molecules able to take advantage

of lipophilic interactions and to interact specifically with the secondary hydrophilic regions. Further docking experiments suggested that favorable interactions could also be established by positively charged groups in the ligands and the negatively charged regions of the protein created by Asp 367 and Glu 358 side chains. Following this analysis, candidates for the R group in **4** were selected among commercially available carboxylic acids featuring aromatic groups and/or hydroxyl groups, amino groups or acetamides.

The initial set of compounds were synthesized starting from amine **5**,<sup>14</sup> which in turn was obtained from tri-*O*-acetyl-L-fucosylazide **6** and the protected (1*S*,2*R*)- $\beta$ -amino acid **7**, as we have previously described (Scheme 1).<sup>14</sup>

The coupling reactions between amine **5** and the acid partners ( $\text{RCO}_2\text{H}$ , Scheme 2) were performed using either HBTU, acid chloride or EDC/HOBt activation, as described in the Supplementary Information. These conditions afforded the protected ligands **8** in variable yields (between 43 and 80%) after isolation by solid phase extraction and chromatographic purification. Removal of the protecting groups under standard Zemplén's conditions gave the required compounds **3**. If the acid partner carried a Boc protection, this was first removed using a mixture of TFA/ $\text{CH}_2\text{Cl}_2$  (1/5). In this way *ca.* 30 different compounds were prepared (see Supplementary Information for the structure and characterization of the entire library). Scheme 2 shows the structure of those that will be used in the following discussion.



**Scheme 2** Coupling reaction of amine **5** with acid partners.

DC-SIGN affinity for the entire set of 30 compounds of general formula **3** was estimated using a surface plasmon resonance (SPR) biosensor in a competition assay which we have previously described.<sup>14</sup> The assay allows an affinity evaluation of all ligands relative to one another on the basis of their percentage inhibition

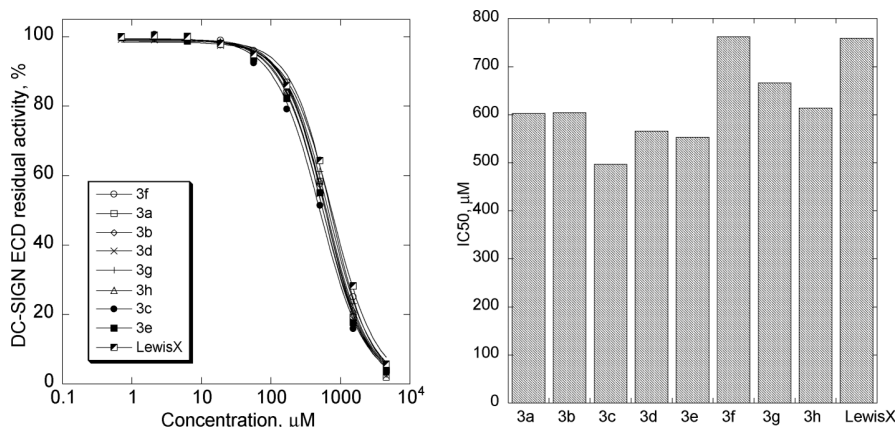


Fig. 4 The dependency of DC-SIGN ECD percent activity on concentrations of corresponding compounds. (see original sensorgrams in SI-Fig.4).

of DC-SIGN binding to immobilized mannosylated bovine serum albumin (Man-BSA). The commercially available Man-BSA used in these assays contain an average of 12 glycosylation sites displaying the Man $\alpha$ 1-3[Man $\alpha$ 1-6]Man branched trisaccharide. Man-BSA was covalently attached to a carboxymethyl dextran-functionalized gold SPR sensor chip CM4. Inhibition studies were then performed using extracellular domain (ECD) of DC-SIGN (20  $\mu$ M) injected alone or in the presence of a constant concentration (300  $\mu$ M) of the ligands. At this concentration, for the particular chip used in the assay, Le<sup>x</sup> exhibited 25% of inhibition. All the molecules **3** synthesized showed a similar efficiency, independent of the nature of the R group, and none improved significantly over the activity of **3a**. To confirm these data, complete inhibition curves were obtained and IC<sub>50</sub> values were estimated for the selected group of compounds **3a–h** shown in Scheme 2. The results were totally consistent with the previous observation (Fig. 4).

Interestingly, when an analogous group of compounds **9** (Fig. 5), obtained by reaction of fucosylazide **6** with  $\beta$ -alanine rather than with **7**, was examined in SPR single point assays at 300  $\mu$ M concentration, similar inhibition values (25–30%) were obtained. Thus, these simple  $\alpha$ -fucosyl- $\beta$ -alanyl amides showed a similar affinity for DC-SIGN as Le<sup>x</sup> and all the compounds **3** synthesized.

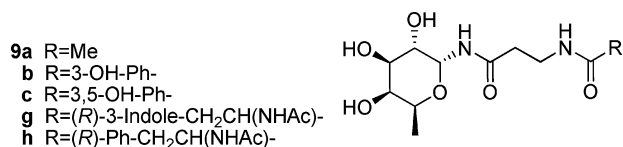


Fig. 5  $\beta$ -Alanine derivatives **9**.

These observations, that are in striking contrast with the expectations derived from docking studies, strongly suggest that the R substituent in **3** is not reaching the secondary binding site identified by the docking algorithm and may not be interacting at all with the protein. The unexpected results obtained with  $\alpha$ -fucosyl- $\beta$ -alanyl amides **9**, whilst providing us with very simple ligands of high efficiency,<sup>22</sup> confirm that the (1*S*,2*R*)-2-amino-cyclohexanecarboxylic acid scaffold selected for the synthesis of **3** does not enforce optimal interaction of the secondary residue with the protein. NMR studies are currently in progress to assess the

structural details of the interaction of DC-SIGN with compounds **3**.<sup>23</sup> To further explore the role of the  $\beta$ -amino acid structure in defining ligand–protein interaction, the configuration of the scaffold was varied systematically and a third set of compounds was synthesized, where the R fragment was kept unchanged and the  $\beta$ -configuration was systematically permutated.

The new set of compounds **10–12** (Fig. 6) were synthesized starting from  $\beta$ -amino acids **13–15** (Fig. 6) using the synthetic sequence employed for **3** and shown in Schemes 1 and 2. The syntheses of the enantiomerically pure isomeric amino acids **13–15** and of the corresponding fucosylamide derivatives **10–12** are described in the Supplementary Information. The IC<sub>50</sub> obtained by SPR analysis of **10–12** are collected in Fig. 7 and compared to selected data obtained for **3**, for Le<sup>x</sup> and for L-fucose (see SI-Fig 5 for original sensorgrams).

The IC<sub>50</sub> values obtained for **3a** and **3b** are consistent with previous measurements (Fig. 4). The data confirm that the activity of most fucosylamides is close to that of Le<sup>x</sup>. The series of compounds **10a–d** show a larger increase of the affinity (a factor of 3) on passing from the acetamide **10a** (R = Me) to the aromatic amides **10b–d**, suggesting a possible role of the aromatic group in the interaction with the protein. This series, which is built on the (1*R*,2*S*)-2-amino-cyclohexanecarboxylic acid scaffold **13**, also contains the strongest ligands so far, the hydroxybenzoic acid derivatives **10b** and **10c** (IC<sub>50</sub> 470  $\mu$ M) and is therefore the best current candidate for further optimization.

To analyze the selectivity properties of the compounds screened, we developed an additional SPR analysis for Langerin binding properties. As for DC-SIGN, the ability of Langerin to bind to a surface functionalized with Man-BSA was tested. In the case of Langerin, binding to Man-BSA as well as to the dextran matrix (SI-Fig 1) was observed. Therefore, the dextran/Man-BSA surface was considered as a combined ligand of Langerin ECD. Upon titration of the surface with Langerin, a similar saturation curve than for DC-SIGN was obtained (SI-Fig. 2 and 3). Langerin displayed an apparent *K*<sub>d</sub> for this surface of 10.3  $\mu$ M. Indeed, the two lectins exhibited comparable affinity for this surface, thus the same fixed concentration of Langerin was used in the SPR-based competition assay, which allowed a direct comparison of the binding inhibition properties of the compounds for Langerin *versus* DC-SIGN.

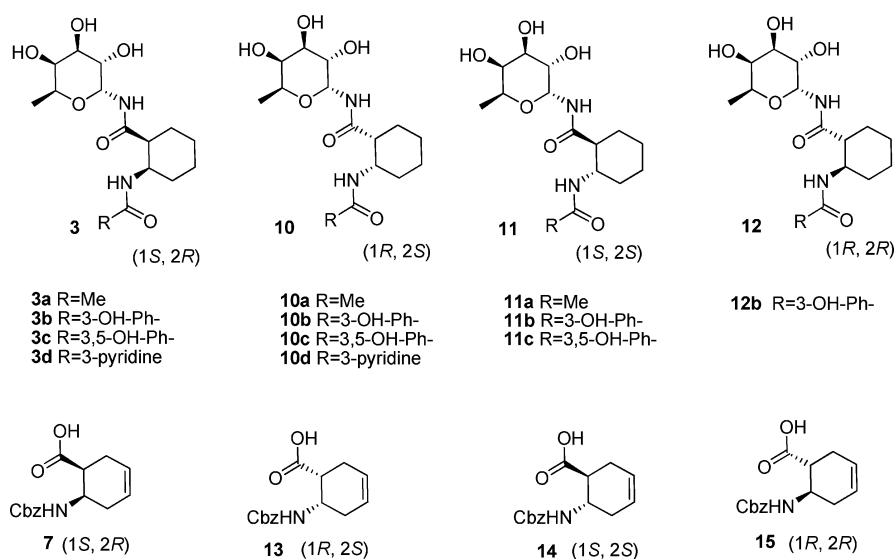


Fig. 6 Library of compounds with different stereochemistry in the cyclohexane scaffold.

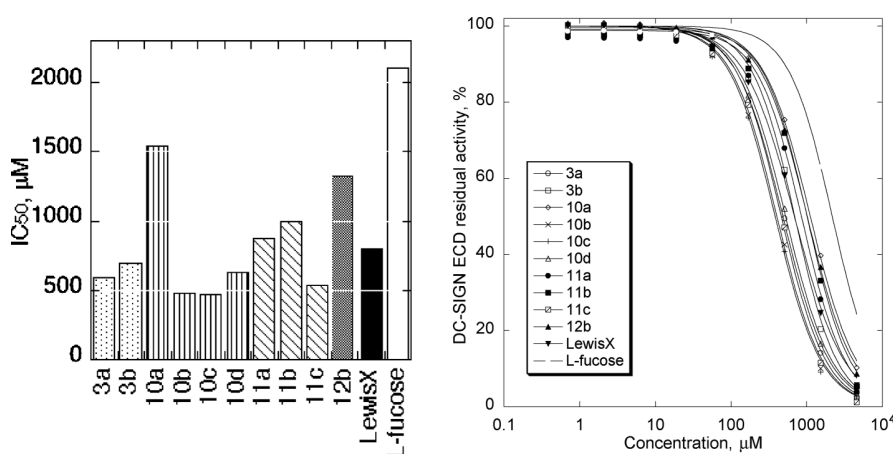


Fig. 7 The  $IC_{50}$  values obtained for compounds 10–12 by SPR inhibition assay.  $IC_{50}$  of L-fucose, Le<sup>x</sup>, 3a and 3b measured in the same conditions are shown for comparison.

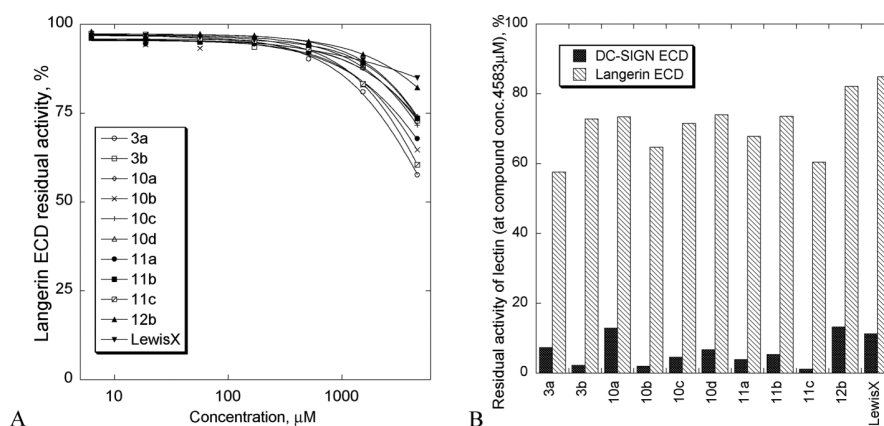
Compounds initially tested for their DC-SIGN inhibitory potency (Fig. 7) were evaluated with Langerin ECD. The results are shown in Fig. 8A (see also SI-Fig.6). The inhibitory potency of the fucosylated mimics is so low against Langerin that it was not possible to determine an  $IC_{50}$ . A crude comparison of the inhibitors' properties towards DC-SIGN and Langerin was obtained by comparing the residual activity of both lectins at the highest concentration tested for each compound (Fig. 8B).

The data show that many of the fucosylamides tested display a larger DC-SIGN selectivity than Le<sup>x</sup> and confirm 10b as one of the most interesting elements of this group, both for its DC-SIGN affinity and for its specificity.

## Conclusions

In this paper, we have presented a new library of fucose-based ligands of DC-SIGN, all characterized by the presence of a  $\beta$ -amino acid tether and of a fucosylamide anchor, able to direct the

molecules to the DC-SIGN CRD binding site. We were able to identify many compounds that, compared to the natural ligand Le<sup>x</sup> and its previously reported mimic 2, display a similar DC-SIGN inhibition efficiency at a fraction of the synthetic cost. In particular,  $\alpha$ -fucosylamides 9, derived from  $\beta$ -alanine, are interesting candidates for polyvalent presentations<sup>8d-k</sup> due to their high synthetic accessibility and good ligand efficiency.<sup>22</sup> A second group of compounds,  $\alpha$ -fucosylamides 10 derived from (1*R*,2*S*)-2-amino-cyclohexanecarboxylic acid 13, were also of interest because they yielded the most active and selective ligand of this group (10b). Indeed, it may be important that molecules directed to block the action of DC-SIGN do not interfere with the action of other lectins. We have recently shown that  $\alpha$ -N-fucosylamides of general formula 3 and 9 interact strongly with the L-fucose binding lectin PA-IIL of *Pseudomonas aeruginosa*.<sup>24</sup> In particular, to develop inhibitors of sexually transmitted HIV infections it may be necessary to select DC-SIGN antagonists that do not interfere with the action of Langerin. Like DC-SIGN, Langerin is a



**Fig. 8** A) Inhibition of Langerin ECD binding to Man-BSA immobilized on the dextran surface. B) Residual lectin activity at the highest tested concentration of the ligands (4.6 mM). Langerin, white bars; DC-SIGN, black bars.

membrane C-type lectin known to bind to HIV-1. However, whilst interaction with DC-SIGN is used by the virus to invade the host immune system, Langerin is suggested to have protective effects against HIV infection.<sup>16,25</sup> Indeed, Langerhans cells, which are the first dendritic cells to encounter HIV *via* genital mucosa, have been described as a natural barrier for HIV-1 transmission, which is dependent on Langerin expression.<sup>15</sup> The importance of Langerin in HIV protection has been again emphasized in the context of HIV-1/Herpes Simplex Virus type II (HSV-2) co-infection.<sup>26</sup> In this last case, it has been demonstrated that HIV susceptibility of Langerhans cells, and the subsequent virus transmission, could be promoted by HSV-2-dependent abrogation of Langerin's functions. Conversely, DC-SIGN has a well-established role in dendritic cells-mediated HIV-1 transmission.<sup>6</sup> Thus, the selectivity of the ligands for DC-SIGN relative to Langerin was also tested. Indeed, we described here for the first time, simultaneous screening of artificial compounds towards both lectins. In agreement with literature data on the natural ligands of these lectins, a low inhibitory potency of Le<sup>x</sup> towards Langerin has been observed in contrast to DC-SIGN. The capacity of Le<sup>x</sup> and of the other fucosylated derivatives to inhibit Langerin binding is so low that we could not perform a full inhibition curve in a reasonable range of concentration. Therefore, selectivity of the fucose-based ligands tested was assessed by comparing residual lectin activity at the highest ligand concentration tested. Most of the  $\alpha$ -fucosylamides assayed were found to be more DC-SIGN selective than the natural ligand Le<sup>x</sup> and therefore they are more likely to be turned into therapeutically useful tools against sexually transmitted HIV infection.<sup>8</sup>

## Experimental

### Langerin and DC-SIGN ECD expression and purification

Langerin ECD constructs (comprising residue 68–328) have been overexpressed using a pET30b derived vector in BL21(DE3) as described previously.<sup>27</sup> The protein was expressed as inclusion body, refolded and purified to homogeneity in a functional form as already described.<sup>22</sup> DC-SIGN ECD protein (residue 66–404) has been overexpressed and purified as described previously.<sup>5a</sup>

### Surface plasmon resonance analysis

All experiments were performed on a Biacore 3000 using functionalized CM4 sensor chips and the corresponding reagents from Biacore. Two flow cells were activated as previously described.<sup>28</sup> Flow cell one was then blocked with 30  $\mu\text{L}$  of 1 M ethanolamine and used as a control surface. The second one was treated with BSA-Man $\alpha$ 1-3[Man $\alpha$ 1-6]Man (Man-BSA, Dextra) (60  $\mu\text{g mL}^{-1}$ ) in 10 mM acetate buffer, pH 4. Remaining activated groups were blocked with 30  $\mu\text{L}$  of 1 M ethanolamine. The final density immobilized on the surface of the second flow cell was 2000 RU. The Man-BSA used to functionalize CM4 chip harbours 12 glycosylation sites according to manufacturer.

Two types of SPR tests were set up for the evaluation of glycomimic compounds. The single point inhibition assay was used for fast screening of compound selectivity. Here, either DC-SIGN or Langerin at concentration of 20  $\mu\text{M}$  were incubated with corresponding compounds (300  $\mu\text{M}$  final concentration) and 20  $\mu\text{L}$  of the samples were co-injected over Man-BSA surface. The lectin steady state binding responses were extracted from the sensorgrams and compared with the responses of compound-free lectin injections and converted to inhibition percent values. In the case of Langerin inhibition assay, even with parallel functionalization of the reference surface with non-glycosylated BSA, some interaction of Langerin with the dextran matrix still remains. Indeed, the dextran/Man-BSA surface has been considered as a combined ligand of Langerin ECD (see Supplementary Information for more details).

The second type of SPR test was used to estimate the relative compound affinity to the lectins on the basis of their IC<sub>50</sub> values. This was accomplished in the same manner as in the single point inhibition assay, except that both lectins were incubated with increasing concentrations (from 0 to 5000  $\mu\text{M}$ ) of the corresponding compounds, the injected sample volumes were 13  $\mu\text{L}$ , and the resulting binding responses were converted to residual lectin activity values, which were plotted against concentration values of the compounds. The relative IC<sub>50</sub> values for each compound were determined by fitting four parameter logistic model (eqn (1)) to the experimental data.

$$y = \text{bot} + \frac{\text{top} - \text{bot}}{1 + \left( \frac{x}{\text{IC}_{50}} \right)^{\text{slope}}} \quad (1)$$

where  $y$  is the percent activity,  $x$  is the corresponding concentration, bot and top are the lowest and the highest values of percent activity, respectively.

In both type of experiments 5  $\mu\text{l}$  flow rate was used and the running buffer was 25 mM Tris-HCl, pH 8, 150 mM NaCl, 4 mM CaCl<sub>2</sub>, and 0.005% of P20 surfactant. All the samples were prepared in the running buffer.

The stability of the surface during a campaign was evaluated by DC-SIGN ECD binding capacity as a function of the number of cycles. The chip surfaces demonstrated a strong stability with negligible decrease of the binding capacity of only 0.06% per cycle (see data in the Supplementary Information).

## Modeling

**Grid analysis.** The properties of the protein surface in the vicinity of the Ca<sup>2+</sup> site were determined using the protein crystal structure derived from the complex DC-SIGN-Le<sup>x</sup> (pdb code 1SL5<sup>12</sup>) and the program GRID<sup>17</sup> (version 22). In particular the DRY probe and the WATER probe were used to identify hydrophilic and lipophilic regions of the binding site. The accompanying program GREAT was used to check the crystal structure file and to prepare the file of coordinates in standard PDB format, which is used as an input for the program GRIN. This program, which prepares the input for the main program GRID, was used to automatically assign atom types and charges for every atom of the protein, using provided standard parameters. Calculations of the interaction energy between the probe and each atom of the protein were performed on a box (36.6 Å × 23.0 Å × 16.6 Å per side) centered on the protein, with a grid spacing of 0.2 Å (NPLA = 5) and its value was evaluated at each grid point. A dielectric constant of 80 was used to simulate a bulk aqueous phase, while a dielectric constant of 4 was assigned to the interior of the protein.

The output, which consists of an array of interaction energies, can be visualized as contour surfaces at appropriate energy levels together with the protein structure. Contours at negative energy levels delineate regions of attraction between probe and protein, whereas positive energy levels define the surface of the protein. Visual inspection of the contour surfaces superimposed on the active site of DC-SIGN enabled the identification of the most favored hydrophilic and lipophilic regions, facilitating the interpretation of protein–ligand interaction. Moreover interaction energy values between the WATER probe and the protein were used to identify the important structural water molecules out of all the crystallized water molecules found in the X-ray structure. An arbitrary cut-off of −10.42 kcal mol<sup>−1</sup> (the most negative energy value was −14.97 kcal mol<sup>−1</sup>) was chosen to detect the most favorable hot spots for a water molecule. This cut off allowed us to identify two important structural water molecules, corresponding to W13 and W36 in the crystal structure of DC-SIGN-Le<sup>x</sup> complex (1SL5). The DC-SIGN crystal structure including these two water molecules was subsequently used for additional docking runs.

**Docking calculations.** Docking calculations were performed using the program GLIDE 4.5.<sup>20</sup> The initial setup for the receptor preparation before docking runs was performed using Schrödinger's 'Protein Preparation Wizard', starting from the X-ray crystal structure of the DC-SIGN-Le<sup>x</sup> complex. All crystallographic water molecules, except for W13 and W36, were deleted, bond orders assigned and hydrogen atoms added. The assignments of protonation states for basic and acidic residues were based on the optimization of hydrogen bonding patterns. The final minimization of the complex was not performed with the Preparation Wizard default, but the complex was minimized (500 steps, Truncated Newton Conjugate Gradient method) in implicit water (GB/SA<sup>21</sup> model) using MacroModel<sup>29</sup> with the AMBER\* force field (dielectric constant 1, cut off extended, convergence on gradient with threshold of 0.05). The oxygen atoms of W13 and W36 as well as the Ca<sup>2+</sup> ion were anchored to the original position through a harmonic potential during minimization. At the end of the minimization, the root-mean-square deviation (RMSD) of all heavy atoms was within 0.34 Å of the crystallographic positions. Docking calculations were performed in Standard Precision mode with standard OPLS-AA(2001)<sup>30</sup> force field; non-planar conformations of amide bonds were penalized, Van der Waals radii were scaled by 0.80 and the partial charge cut off was fixed to 0.15. The shape and properties of the binding site were mapped onto grids with dimensions of 36 Å (enclosing box) and 14 Å (ligand diameter midpoint box), centered on the ligand in the X-ray structure of the DC-SIGN-Le<sup>x</sup> complex. Docking was constrained in the Ca<sup>2+</sup> binding site by specifying a reference core corresponding to the C1–C6 carbon and O3–O4–O5 oxygen atoms of the fucose residue of the reference ligand in the X-ray structure of the DC-SIGN-Le<sup>x</sup> complex: ligands that feature the same core moiety as the reference ligand are subject to the constraint. The RMSD tolerance for the position of the core was set to 3.5 Å. This parameter enforces the fucose moiety to be located within 3.5 Å of the fucose residue in the reference ligand: ligand poses that do not match this constraint are screened out. These constraints allow coordination of the Ca<sup>2+</sup> ion by the hydroxy groups OH-2 and OH-3 of fucose, as observed in several crystal structure of C-type lectins complexed with fucose-containing ligands. The RMSD tolerance of 3.5 Å was selected to allow docking poses to explore both possible binding modes of the vicinal diol. Indeed, crystallographic and NMR data on DC-SIGN show that the monosaccharide moiety of ligands can coordinate the Ca<sup>2+</sup> ion with the vicinal hydroxyl groups in two possible orientations, differing by a 180° rotation. The quality of this docking protocol was validated by re-docking the Le<sup>x</sup> ligand in the DC-SIGN-Le<sup>x</sup> complex, which yielded a ligand pose that could be superimposed with crystalline Le<sup>x</sup> with RMSD of 1.71 Å (0.31 Å for the fucose residue).

## Synthesis

Solvents were dried by standard procedures: dichloromethane, methanol, *N,N*-diisopropylethylamine and triethylamine were dried over calcium hydride, chloroform and pyridine were dried over activated molecular sieves. Reactions requiring anhydrous conditions were performed under nitrogen. <sup>1</sup>H, <sup>13</sup>C and <sup>31</sup>P-NMR spectra were recorded at 400 MHz on a Bruker AVANCE-400 instrument. Chemical shifts ( $\delta$ ) for <sup>1</sup>H and <sup>13</sup>C spectra are

expressed in ppm relative to internal Me<sub>4</sub>Si as standard. Signals are abbreviated as s, singlet; bs, broad singlet; d, doublet; t, triplet; q, quartet; m, multiplet. Mass spectra were obtained with a Bruker ion-trap Esquire 3000 apparatus (ESI ionization) or an Autospec Fission Instrument (FAB ionization). HRMS (FT-ICR, ESI) were obtained with an Apex II instrument. Thin layer chromatography (TLC) was carried out with pre-coated Merck F<sub>254</sub> silica gel plates. Flash chromatography (FC) was carried out with Macherey-Nagel silica gel 60 (230–400 mesh). The libraries were synthesized through the common approach shown in Scheme 1.<sup>14</sup> All synthetic schemes and procedures, the synthesis and characterization of new intermediates, the synthesis and characterization of the full library of ligands (**3**, **9**, **11** and **12**) as well as their <sup>1</sup>H- and <sup>13</sup>C-NMR spectra are collected in the Supplementary Information. Below we report the characterization of compounds **10a–d** (from (1*R*,2*S*)-2-aminocyclohexanecarboxylic acid). Compounds **3a** and **5** were previously described.<sup>14</sup>

***N*-[(1*R*,2*S*)-2-Acetamido-cyclohexanecarboxyl]- $\alpha$ -L-fucopyranosylamine (**10a**)**

The crude hydrogenation product of **28** (see Supplementary Information–SI-Scheme 3) was used in the general acetylation method (see Supplementary Information). The product was purified by flash chromatography on silica gel (AcOEt, *R*<sub>f</sub> 0.29). Yield: 27 mg (82%). Zemplen deprotection and flash chromatography (85:15 CHCl<sub>3</sub>:MeOH, *R*<sub>f</sub> 0.18) afforded **10a**. Yield: 13 mg (65%). <sup>1</sup>H-NMR (400 MHz, CD<sub>3</sub>OD):  $\delta$  (ppm) = 1.18 (d, 3H, *J*<sub>5-6</sub> = 6.5 Hz, H<sub>F6</sub>), 1.32–1.50 (m, 2H, H<sub>Cy4ax</sub> and H<sub>Cy5ax</sub>), 1.55–1.68 (m, 3H, H<sub>Cy6ax</sub>, H<sub>C3ax</sub> and H<sub>C4eq</sub> or H<sub>C5eq</sub>), 1.73–1.78 (m, 1H, H<sub>Cy4eq</sub> or H<sub>Cy5eq</sub>), 1.83–1.92 (m, 1H, H<sub>Cy3eq</sub> and H<sub>C6eq</sub>), 1.94 (s, 3H, Ac-Me), 2.71–2.75 (m, 1H, H<sub>Cy1</sub>), 3.63–3.66 (m, 1H, H<sub>F4</sub>), 3.74 (dd, *J*<sub>3-4</sub> = 3.3 Hz, *J*<sub>3-2</sub> = 10.3 Hz, 1H, H<sub>F3</sub>), 3.79 (q, *J*<sub>5-6</sub> = 6.5 Hz, 1H, H<sub>F5</sub>), 3.94 (dd, *J*<sub>1-2</sub> = 5.6 Hz, *J*<sub>2-3</sub> = 10.2 Hz, 1H, H<sub>F2</sub>), 4.18–4.25 (m, 1H, H<sub>Cy2</sub>), 5.51 (d, 1H, *J*<sub>1-2</sub> = 5.6 Hz, H<sub>F1</sub>). <sup>13</sup>C-NMR (100 MHz, CD<sub>3</sub>OD):  $\delta$  (ppm) = 17.0 (C<sub>F6</sub>), 23.0 (Ac-Me), 23.6 (C<sub>Cy4</sub> or C<sub>Cy5</sub>), 24.2 (C<sub>Cy4</sub> or C<sub>Cy5</sub>), 26.8 (C<sub>Cy6</sub>), 30.7 (C<sub>Cy3</sub>), 40.3 (C<sub>Cy1</sub>), 46.4 (C<sub>Cy2</sub>), 68.3 (C<sub>F2</sub>), 68.8 (C<sub>F5</sub>), 71.8 (C<sub>F3</sub>), 73.3 (C<sub>F4</sub>), 78.3 (C<sub>F1</sub>), 173.1 (NHCO), 177.5 (NHCO). HRMS (FT-ICR, ESI): *m/z* calcd for C<sub>15</sub>H<sub>26</sub>N<sub>2</sub>O<sub>6</sub>: 353.16831 [M + Na]<sup>+</sup>; found: 353.16838. [ $\alpha$ ]<sub>D</sub> –77.5 (c 0.35, EtOH)

***N*-[(1*R*,2*S*)-2-(3-Hydroxybenzamido)cyclohexanecarboxyl]- $\alpha$ -L-fucopyranosylamine (**10b**)**

The crude hydrogenation product of **28** (see Supplementary Information–SI-Scheme 3) was coupled with 3-hydroxybenzoic acid using the HBTU general procedure (see Supplementary Information) and the product was purified by flash chromatography (6:4 AcOEt:petroleum ether, *R*<sub>f</sub> 0.38). Yield: 13 mg (31%). Zemplen deprotection and flash chromatography (85:15 chloroform:methanol, *R*<sub>f</sub> 0.17) afforded **10b**. Yield: 8 mg (89%). <sup>1</sup>H-NMR (400 MHz, CD<sub>3</sub>OD):  $\delta$  (ppm) = 1.18 (d, 3H, *J*<sub>5-6</sub> = 6.5 Hz, H<sub>F6</sub>), 1.42–1.55 (m, 2H, H<sub>Cy4ax</sub> and H<sub>Cy5ax</sub>), 1.61–1.76 (m, 4H, H<sub>Cy6ax</sub>, H<sub>Cy3ax</sub>, H<sub>Cy4eq</sub> and H<sub>Cy5eq</sub>), 1.94–2.02 (m, 1H, H<sub>Cy6eq</sub>), 2.16–2.24 (m, 1H, H<sub>Cy3eq</sub>), 2.86–2.91 (m, 1H, H<sub>Cy1</sub>), 3.63–3.66 (m, 1H, H<sub>F4</sub>), 3.74–3.80 (m, 2H, H<sub>F5</sub> and H<sub>F3</sub>), 3.91–3.95 (m, 1H, H<sub>F2</sub>), 4.28–4.33 (m, 1H, H<sub>Cy2</sub>), 5.52–5.54 (m, 1H, H<sub>F1</sub>) 6.90–6.95 (m, 1H, Ar), 7.18–7.28 (m, 3H, Ar). <sup>13</sup>C-NMR (100 MHz, CD<sub>3</sub>OD):  $\delta$  (ppm) = 17.0 (C<sub>F6</sub>),

23.8 (C<sub>Cy4</sub> or C<sub>Cy5</sub>), 24.2 (C<sub>Cy4</sub> or C<sub>Cy5</sub>), 27.8 (C<sub>Cy6</sub>), 30.4 (C<sub>Cy3</sub>), 46.2 (C<sub>Cy1</sub>), 50.2 (C<sub>Cy2</sub>), 68.2 (C<sub>F2</sub>), 68.8 (C<sub>F5</sub>), 71.7 (C<sub>F3</sub>), 73.3 (C<sub>F4</sub>), 78.4 (C<sub>F1</sub>), 115.3 (Ar), 119.4 (Ar), 119.7 (Ar), 130.8 (Ar), 137.5 (C<sub>quart.Ar</sub>), 158.9 (C<sub>quart.Ar</sub>), 170.0 (NHCO), 178.0 (NHCO). HRMS (FT-ICR, ESI): *m/z* calcd for C<sub>20</sub>H<sub>28</sub>N<sub>2</sub>O<sub>7</sub>: 431.17887 [M + Na]<sup>+</sup>; found: 431.17948. [ $\alpha$ ]<sub>D</sub> –53.1 (c 0.25, MeOH)

***N*-[(1*R*,2*S*)-2-(3,5-Dihydroxybenzamido)cyclohexanecarboxyl]- $\alpha$ -L-fucopyranosylamine (**10c**)**

The crude hydrogenation product of **28** (see Supplementary Information–SI-Scheme 3) was coupled with 3,5-dihydroxybenzoic acid using the HBTU general procedure (see Supplementary Information) and the product was purified by flash chromatography (7:3 ethyl acetate:n-hexane, *R*<sub>f</sub> 0.43). Yield: 17 mg (37%). Zemplen deprotection and flash chromatography (85:15 chloroform:methanol, *R*<sub>f</sub> 0.14) afforded **10c**. Yield: 10 mg (91%). <sup>1</sup>H-NMR (400 MHz, CD<sub>3</sub>OD):  $\delta$  (ppm) = 1.19 (d, 3H, *J*<sub>5-6</sub> = 6.5 Hz, H<sub>F6</sub>), 1.42–1.55 (m, 2H, H<sub>Cy4ax</sub> and H<sub>Cy5ax</sub>), 1.61–1.73 (m, 4H, H<sub>C6ax</sub>, H<sub>Cy3ax</sub>, H<sub>Cy4eq</sub> and H<sub>Cy5eq</sub>), 1.92–2.02 (m, 1H, H<sub>Cy6eq</sub>), 2.16–2.22 (m, 1H, H<sub>Cy3eq</sub>), 2.85–2.90 (m, 1H, H<sub>Cy1</sub>), 3.64–3.66 (m, 1H, H<sub>F4</sub>), 3.77 (dd, *J*<sub>3-4</sub> = 3.3 Hz, *J*<sub>2-3</sub> = 10.3 Hz, 1H, H<sub>F3</sub>), 3.79 (q, *J*<sub>5-6</sub> = 6.5 Hz, 1H, H<sub>F5</sub>), 3.95 (dd, 1H, *J*<sub>1-2</sub> = 5.6 Hz, *J*<sub>2-3</sub> = 10.3 Hz, H<sub>F2</sub>), 4.25–4.30 (m, 1H, H<sub>Cy2</sub>), 5.55 (d, 1H, *J*<sub>1-2</sub> = 5.6 Hz, H<sub>F1</sub>) 6.39 (bs, 1H, Ar), 6.68 (bs, 2H, Ar). <sup>13</sup>C-NMR (100 MHz, CD<sub>3</sub>OD):  $\delta$  (ppm) = 17.0 (C<sub>F6</sub>), 23.8 (C<sub>Cy4</sub> or C<sub>Cy5</sub>), 24.1 (C<sub>Cy4</sub> or C<sub>Cy5</sub>), 27.9 (C<sub>Cy6</sub>), 30.4 (C<sub>Cy3</sub>), 46.2 (C<sub>Cy1</sub>), 50.2 (C<sub>Cy2</sub>), 68.2 (C<sub>F2</sub>), 68.8 (C<sub>F3</sub> or C<sub>F5</sub>), 71.7 (C<sub>F3</sub> or C<sub>F5</sub>), 73.3 (C<sub>F4</sub>), 78.4 (C<sub>F1</sub>), 106.8 (Ar), 106.8 (Ar), 138.1 (C<sub>quart.</sub>, Ar.), 160.0 (C<sub>quart.</sub>, Ar.), 170.0 (NHCO), 178.0 (NHCO). HRMS (FT-ICR, ESI): *m/z* calcd for C<sub>20</sub>H<sub>28</sub>N<sub>2</sub>O<sub>8</sub>: 447.17379 [M + Na]<sup>+</sup>; found: 447.17407. [ $\alpha$ ]<sub>D</sub> –65.7 (c 0.20, MeOH)

***N*-[(1*R*,2*S*)-2-(3-Pyridinecarboxamido)cyclohexanecarboxyl]- $\alpha$ -L-fucopyranosylamine (**10d**)**

The crude hydrogenation product of **28** (see Supplementary Information–SI-Scheme 3) was coupled with nicotinic acid using the HBTU general procedure (see Supplementary Information) and the product was purified by flash chromatography (97:3 ethyl acetate:triethyl amine, *R*<sub>f</sub> 0.17). Yield: 19 mg (50%) Zemplen deprotection and flash chromatography (85:15 chloroform:methanol, *R*<sub>f</sub> 0.17) afforded **10d**. Yield: 6 mg (43%). <sup>1</sup>H-NMR (400 MHz, CD<sub>3</sub>OD):  $\delta$  (ppm) = 1.19 (d, 3H, *J*<sub>5-6</sub> = 6.5 Hz, H<sub>F6</sub>), 1.43–1.55 (m, 2H, H<sub>Cy4ax</sub> and H<sub>Cy5ax</sub>), 1.67–1.78 (m, 4H, H<sub>Cy6ax</sub>, H<sub>Cy3ax</sub>, H<sub>Cy4eq</sub> and H<sub>C5eq</sub>), 1.94–2.04 (m, 1H, H<sub>Cy3eq</sub>), 2.14–2.21 (m, 1H, H<sub>Cy6eq</sub>), 2.88–2.92 (m, 1H, H<sub>Cy1</sub>), 3.63–3.66 (m, 1H, H<sub>F4</sub>), 3.74 (dd, *J*<sub>3-4</sub> = 3.4 Hz, *J*<sub>2-3</sub> = 10.3 Hz, 1H, H<sub>F3</sub>), 3.78 (q, *J*<sub>5-6</sub> = 6.5 Hz, 1H, H<sub>F5</sub>), 3.93 (dd, 1H, *J*<sub>1-2</sub> = 5.6 Hz, *J*<sub>2-3</sub> = 10.2 Hz, H<sub>F2</sub>), 4.37–4.42 (m, 1H, H<sub>Cy2</sub>), 5.53 (d, 1H, *J*<sub>1-2</sub> = 5.6 Hz, H<sub>F1</sub>), 7.50–7.53 (m, 1H, H<sub>Ar5</sub>), 8.20 (d, 1H, *J*<sub>Ar5-Ar6</sub> = 8.0 Hz, H<sub>Ar6</sub>), 8.65 (d, 1H, *J*<sub>Ar4-Ar5</sub> = 4.5 Hz, H<sub>Ar4</sub>), 8.92 (s, 1H, H<sub>Ar2</sub>). <sup>13</sup>C-NMR (100 MHz, CD<sub>3</sub>OD):  $\delta$  (ppm) = 17.0 (C<sub>F6</sub>), 23.7 (C<sub>Cy4</sub> and C<sub>Cy5</sub>), 24.1 (C<sub>Cy3</sub>), 27.4 (C<sub>Cy6</sub>), 30.4 (C<sub>Cy1</sub>), 50.4 (C<sub>Cy2</sub>), 68.2 (C<sub>F2</sub>), 68.8 (C<sub>F5</sub>), 71.7 (C<sub>F3</sub>), 73.3 (C<sub>F4</sub>), 78.3 (C<sub>F1</sub>), 125.1 (C<sub>Ar5</sub>), 132.7 (C<sub>quart.Ar</sub>), 137.4 (C<sub>Ar6</sub>), 149.5 (C<sub>Ar2</sub>), 152.5 (C<sub>Ar4</sub>), 167.95 (NHCO), 177.76 (NHCO). HRMS (FT-ICR, ESI): *m/z* calcd for C<sub>19</sub>H<sub>27</sub>N<sub>3</sub>O<sub>6</sub>: 416.17921 [M + Na]<sup>+</sup>; found: 416.17934. [ $\alpha$ ]<sub>D</sub> –81.9 (c 0.15, MeOH)



## Acknowledgements

This work was supported by the FIRB program CHEM-PROFARMANET (RBPR05NWWC), the Marie Curie ITN FP7 project CARMUSYS (PITN-GA-2008-213592) and Comune di Milano (Convenzione 55/2008). Exact mass were obtained from CIGA (Centro interdipartimentale grandi apparecchiature, Università degli Studi di Milano).

## References

- (a) J. Banchereau and R. M. Steinman, *Nature*, 1998, **392**, 245–252; (b) D. N. J. Hart, *Blood*, 1997, **90**, 3245–3287; (c) I. Mellman and R. M. Steinman, *Cell*, 2001, **106**, 255–258.
- (a) A. Cambi and C. G. Figdor, *Curr. Opin. Immunol.*, 2005, **17**, 345–351; (b) E. McGreal, J. Miller and S. Gordon, *Curr. Opin. Immunol.*, 2005, **17**, 18–24; (c) C. Théry and S. Amigorena, *Curr. Opin. Immunol.*, 2001, **13**, 45–51.
- C. G. Figdor, Y. van Kooyk and G. J. Adema, *Nat. Rev. Immunol.*, 2002, **2**, 77–84.
- T. B. H. Geijtenbeek, R. Torensma, S. J. Van Vliet, G. C. F. van Duynhoven, G. J. Adema, Y. van Kooyk and C. G. Figdor, *Cell*, 2000, **100**, 575–585.
- (a) G. Tabarani, M. Thepaut, D. Stroebel, C. Ebel, C. Vives, P. Vachette, D. Durand and F. Fieschi, *J. Biol. Chem.*, 2009, **284**, 21229–21240; (b) D. Serrano-Gomez, E. Sierra-Filardi, R. T. Martinez-Nunez, E. Caparros, R. Delgado, M. A. Munoz-Fernandez, M. A. Abad, J. Jiménez-Barbero, M. Leal and A. L. Corbi, *J. Biol. Chem.*, 2008, **283**, 3889–3903; (c) H. Feinberg, Y. Guo, D. A. Mitchell, K. Drickamer and W. I. Weis, *J. Biol. Chem.*, 2005, **280**, 1327–1335; (d) D. A. Mitchell, A. J. Fadden and K. Drickamer, *J. Biol. Chem.*, 2001, **276**, 28939–28945.
- (a) T. B. H. Geijtenbeek, D. S. Kwon, R. Torensma, S. J. Van Vliet, G. C. F. Van Duynhoven, J. Middel, I. L. M. H. A. Cornelissen, H. S. L. M. Nottet, V. N. Kewal Ramani, D. R. Littman, C. G. Figdor and Y. Van Kooyk, *Cell*, 2000, **100**, 587–597; (b) P.-Y. Lozach, H. Lortat-Jacob, A. De Lacroix De Lavalette, I. Staropoli, S. Foug, A. Amara, C. Houles, F. Fieschi, O. Schwartz, J.-L. Virelizier, F. Arenzana-Seisdedos and R. Altmeyer, *J. Biol. Chem.*, 2003, **278**, 20358–20366; C. P. Alvarez, F. Lasala, J. Carrillo, O. Muniz, A. L. Corbi and R. Delgado, *J. Virol.*, 2002, **76**, 6841–6844.
- (a) T. B. H. Geijtenbeek, S. J. van Vliet, E. A. Koppel, M. Sanchez-Hernandez, C. Vandenbroucke-Grauls, B. Appelmek and Y. van Kooyk, *J. Exp. Med.*, 2003, **197**, 7–17; (b) L. Tailleux, O. Schwartz, J. L. Herrmann, E. Pivert, M. Jackson, A. Amara, L. Legres, D. Dreher, L. P. Nicod, J. C. Gluckman, P. H. Lagrange, B. Gicquel and O. Neyrolles, *J. Exp. Med.*, 2003, **197**, 121–127.
- (a) J. J. Reina, S. Sattin, D. Invernizzi, S. Mari, L. Martinez-Prats, G. Tabarani, F. Fieschi, R. Delgado, P. M. Nieto, J. Rojo and A. Bernardi, *ChemMedChem*, 2007, **2**, 1030–1036; (b) D. A. Mitchell, N. A. Jones, S. J. Hunter, J. M. D. Cook, S. F. Jenkinson, M. R. Wormald, R. A. Dwek and G. W. J. Fleet, *Tetrahedron: Asymmetry*, 2007, **18**, 1502–1510; (c) M. J. Borrok and L. L. Kiessling, *J. Am. Chem. Soc.*, 2007, **129**, 12780–12785; (d) F. Lasala, E. Arce, J. R. Otero, J. Rojo and R. Delgado, *Antimicrob. Agents Chemother.*, 2003, **47**, 3970–3972; (e) J. Rojo and R. Delgado, *J. Antimicrob. Chemother.*, 2004, **54**, 579–581; (f) G. Tabarani, J. J. Reina, C. Ebel, C. Vives, H. Lortat-Jacob, J. Rojo and F. Fieschi, *FEBS Lett.*, 2006, **580**, 2402–2408; (g) S.-K. Wang, P.-H. Liang, R. D. Astronomo, T.-L. Hsu, S.-L. Hsieh, D. R. Burton and C.-H. Wong, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 3690–3695; (h) O. Martinez-Avila, L. M. Bedoya, M. Marradi, C. Clavel, J. Alami and S. Penades, *ChemBioChem*, 2009, **10**, 1806–1809; (i) S. Sattin, A. Daggetti, Anna, M. Thepaut, A. Berzi, M. Sanchez-Navarro, G. Tabarani, J. Rojo, F. Fieschi, M. Clerici and A. Bernardi, *ACS Chem. Biol.*, 2010, **5**, 301–312; (j) O. Martinez-Avila, K. Hijazi, M. Marradi, C. Clavel, C. Campion, C. Kelly and S. Penades, *Chem.-Eur. J.*, 2009, **15**, 9874–9888; (k) K. C. A. Garber, K. Wangkanont, E. E. Carlson and L. L. Kiessling, *Chem. Commun.*, 2010, **46**, 6747–6749.
- E. J. Soilleux, *Clin. Sci.*, 2003, **104**, 437–446.
- S. Pöhlmann, F. Baribaud and R. W. Doms, *Trends Immunol.*, 2001, **22**, 643–646.
- (a) S. Pöhlmann, E. J. Soilleux, F. Baribaud, G. J. Leslie, L. S. Morris, J. Trowsdale, B. Lee, N. Cleman and R. W. Doms, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 2670–2675; (b) E. Van Liempt, C. M. C. Bank, P. Mehta, J. J. Garcia-Vallejo, Z. S. Kwar, R. Geyer, R. A. Alvarez, R. D. Cummings, Y. van Kooyk and I. van Die, *FEBS Lett.*, 2006, **580**, 6123–6131; (c) E. Van Liempt, A. Imberty, C. M. C. Bank, S. J. Van Vliet, Y. Van Kooyk, T. B. H. Geijtenbeek and I. Van Die, *J. Biol. Chem.*, 2004, **279**, 33161–33167.
- Y. Guo, H. Feinberg, E. Conroy, D. A. Mitchell, R. Alvarez, O. Blixt, M. E. Taylor, W. I. Weis and K. Drickamer, *Nat. Struct. Mol. Biol.*, 2004, **11**, 591–598.
- H. Feinberg, M. E. Taylor, N. Razi, R. McBride, Y. A. Knirel, S. A. Graham, K. Drickamer and W. I. Weis, *J. Mol. Biol.*, 2011, **405**, 1027–1039.
- G. Timpano, G. Tabarani, M. Anderluh, D. Invernizzi, F. Vasile, D. Potenza, P. M. Nieto, J. Rojo, F. Fieschi and A. Bernardi, *ChemBioChem*, 2008, **9**, 1921–1930.
- For a review on structural studies on the conformation of Lewis oligosaccharides see: E. Yuriev, W. Farrugia, A. M. Scott and P. A. Ramsland, *Immunol. Cell Biol.*, 2005, **83**, 709–717.
- L. de Witte, A. Nabatov, M. Pion, D. Fluittsma, M. A. de Jong, T. de Gruijil, V. Piguet, Y. van Kooyk and T. B. H. Geijtenbeek, *Nat. Med.*, 2007, **13**, 367–371.
- P. J. Goodford, *J. Med. Chem.*, 1985, **28**, 849–857.
- H. Feinberg, D. A. Mitchell, K. Drickamer and W. I. Weis, *Science*, 2001, **294**, 2163–2166.
- H. Feinberg, R. Castelli, K. Drickamer, P. H. Seeberger and W. I. Weis, *J. Biol. Chem.*, 2007, **282**, 4202–4209.
- Glide version 4.5, Schrödinger, LLC NY 2007.
- W. C. Still, A. Tempczyk, R. C. Hawley and T. Hendrickson, *J. Am. Chem. Soc.*, 1990, **112**, 6127–6129.
- A. L. Hopkins, C. R. Groom and A. Alex, *Drug Discovery Today*, 2004, **9**, 430–431.
- C. Guzzi, J. Angulo, F. Doro, J. J. Reina, M. Thépaut, F. Fieschi, A. Bernardi, J. Rojo and P. M. Nieto, submitted.
- M. Andreini, M. Anderluh, A. Audfray, A. Bernardi and A. Imberty, *Carbohydr. Res.*, 2010, **345**, 1400–1407.
- L. de Witte, A. Nabatov and T. B. H. Geijtenbeek, *Trends Mol. Med.*, 2008, **14**, 12–19.
- M. A. W. P. de Jong, L. de Witte, M. Taylor and T. B. H. Geijtenbeek, *J. Immunol.*, 2010, **185**, 1633–1641.
- M. Thépaut, J. Valladeau, A. Nurisso, R. Kahn, B. Arnou, C. Vives, S. Saeland, C. Ebel, C. Monnier, C. Dezutter-Dambuyant, A. Imberty and F. Fieschi, *Biochemistry*, 2009, **48**, 2684–2698.
- F. Halary, A. Amara, H. Lortat-Jacob, M. Messerie, T. Delaunay, C. Houles, F. Fieschi, F. Arenzana-Seisdedos, J.-F. Moreaus and J. Déchanet-Merville, *Immunity*, 2002, **17**, 653–664.
- MacroModel, version 9.5, Schrödinger, LLC, New York, NY, 2007.
- W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225–11235.



## Cinquième partie

### Discussion

---

---

Le prix Nobel 2011 de médecine a été décerné à Bruce A. Beutler, Jules A. Hoffmann et à Ralph M. Steinman pour leurs différentes découvertes sur le système immunitaire. Jules Hoffmann obtient ce titre pour ses travaux sur la caractérisation d'une réponse antifongique chez la drosophile [67]. Il a mis en évidence que l'activation d'une réponse immunitaire après l'infection de drosophiles par des champignons passait par une voie de signalisation Toll, impliquant par conséquent les récepteurs TLRs. Puis Bruce Beutler montra, en 1998, que le récepteur TLR4 était impliqué dans une voie de défense contre les bactéries gram-négative [80]. En effet, il mit en évidence qu'une mutation dans le gène *tlr4* prédispose au développement d'infection par les bactéries gram-négative. Ces deux équipes ont montré le rôle primordial des récepteurs Toll dans la réponse immunitaire innée.

Ralph Steinman partage cette distinction avec Bruce Beutler et Jules Hoffmann pour la découverte des cellules dendritiques et surtout leur implication dans l'activation d'une réponse immunitaire efficace [94, 95]. Il montra le rôle unique de ces cellules dans la capacité d'activation de réponse immunitaire adaptative par la présentation d'antigène aux lymphocytes T. Impliqué dans la compréhension de l'activation de la réponse immunitaire, il s'intéressa aux récepteurs d'agents pathogènes de ces cellules. Il s'orienta bien évidemment sur les TLRs mais aussi sur une autre famille de récepteurs de pathogènes, les CLRs (« C-type Lectin Receptor »). Effectivement les lectines permettent de déclencher des mécanismes immunitaires après la reconnaissance d'agents pathogènes. Par exemple, la reconnaissance de *Pneumocystis carinii* par la lectine Dectine-1 permet d'induire la production d'espèces réactives de l'oxygène [85], ou encore, DC-SIGN induit la défense contre *Mycobacterium tuberculosis* [63] et la Langérine permet d'activer la présentation d'antigènes non-peptidiques dans le cas de *Mycobacterium leprae* [55]. L'étude de cette famille de récepteurs est essentielle à la compréhension de la reconnaissance et de l'induction de mécanismes immunitaires contre des agents pathogènes.

Les études sur la Langérine ont montré différents rôles de cette lectine. Par exemple, de récentes études, orchestrées par Ralph Steinman, montrent la capacité de la Langérine, exprimée dans un sous-type de cellules dendritiques dermiques, à activer des voies de présentation d'antigènes aux lymphocytes T CD4<sup>+</sup> et CD8<sup>+</sup> [56] [43]. Mais, dans les mêmes conditions, les cellules de Langerhans sont incapable de présenter des antigènes *via* la Langérine. La différence d'expression est significative car les cellules dendritiques dermiques Langérine<sup>+</sup> présentent 5% de la quantité retrouvée dans les cellules de Langerhans [43]. Ces études montrent donc des rôles différents de la Langérine dans deux sous-types de cellules dendritiques, les cellules dendritiques dermiques Langérine<sup>+</sup> et les cellules de Langerhans. Ces rôles, parfois contradictoires dans ces deux sous-types cellulaires, sont probablement modulés par une capacité unique de la Langérine à former un organite cellulaire à « elle-

---

seule » dans les LCs, le granule de Birbeck (BG). Il a été montré que la surexpression de cette lectine dans des fibroblastes de souris permettait d'obtenir des structures similaires aux granules de Birbeck des LCs [104]. En 2007, les cellules de Langerhans ont été mises en lumière par leur capacité à reconnaître le VIH, *via* la Langérine, de l'internaliser dans le granule de Birbeck et de l'éliminer [24]. Dans le cas de *Mycobacterium leprae* la Langérine permet d'induire la présentation d'antigène non-peptidique *via* CD1a par les LCs [55]. Cette observation apparaît contradictoire par rapport à l'étude précédente, mais elle est à nuancer du fait de deux modes d'activation différents, par un anticorps d'une part et un pathogène d'autre part. Ces exemples montrent la spécificité des voies induites. CD1a est aussi retrouvé dans les granules de Birbeck [33]. Ceci montre le rôle coopératif de la Langérine et de CD1a dans la présentation d'antigène non-peptidique et donne une piste pour la compréhension du rôle du granule de Birbeck. L'étude de cette lectine permettrait d'aborder la problématique de la reconnaissance de pathogènes par les cellules de Langerhans.

Donc la caractérisation de l'organisation de la Langérine dans les granules de Birbeck est un point fascinant auquel nous nous sommes intéressés dans une première partie.

Une autre particularité de cette lectine, qui a été peu étudiée au début de ce travail, est sa spécificité à l'égard de sucres sulfatés. Nous nous sommes donc focalisés sur la caractérisation de cette spécificité unique, chez les lectines de type-C [44].

Ce travail de thèse a permis de mettre en évidence plusieurs aspects différents de la Langérine dans son implication dans le granule de Birbeck du point de vue structural, et dans les modes de reconnaissance de différents ligands à travers notamment un nouveau mode d'interaction de cette lectine avec les glycosaminoglycanes.

### ***Importance de la Langérine pour la formation des granules de Birbeck***

La Langérine est l'élément indispensable pour la formation de granule de Birbeck. Différentes études montrent qu'elle détermine la formation, la localisation et la structure des granules de Birbeck [104, 105, 107, 100]. La délétion du domaine CRD de la Langérine conduit à l'absence des BGs, par exemple, alors que la délétion du domaine cytoplasmique de cette même lectine affecte l'adressage, la localisation cellulaire et l'organisation des BGs [100]. Le domaine CRD de la Langérine est donc indispensable pour la formation des BGs. Des mutations ponctuelles dans ce domaine montrent des effets drastiques pour la formation et la structure des BGs. En effet, la forme polymorphique humaine de la Langérine, avec la mutation W264R, conduit à la perte des BGs et la mutation F244L dans la Langérine murine conduit à une modification de la structure des BGs [105, 107]. L'étude de ces mutations nous a permis de donner des informations sur l'organisation de la Langérine dans les BGs.

---

En effet, nous avons montré l'importance de la structure tertiaire du domaine CRD de la Langérine, par l'étude de la mutation W264R, et de la structure quaternaire de la Langérine à travers l'étude de la mutation F241L.

Les résultats obtenus nous permettent de mettre en évidence l'importance d'un coeur structural hydrophobe, constitué notamment par le résidu W264, pour le maintien de la structure tertiaire du domaine CRD. La déstabilisation de ce coeur hydrophobe, par la mutation W264R, conduit donc à la perte de la formation des BGs et ainsi montre le rôle indispensable de la structure de ce domaine pour la formation et la structure des BGs.

La mutation F241L, quant à elle, ne conduit pas à des modifications importantes de la fonction et de la structure du domaine CRD de la Langérine. Les modifications observées de la structure cristalline mutante montrent un maintien global de l'intégrité du coeur hydrophobe centré sur le résidu F241. Cependant, la diminution de volume de la chaîne latérale, remplacement du résidu de phénylalanine par un résidu de leucine, entraîne une réorientation de la chaîne latérale du résidu Lys200. La relocalisation de la chaîne latérale du résidu Lys200 conduit à établir une liaison hydrogène avec le résidu Thr245. Ce réarrangement, qui n'a pas d'impact conséquent sur le domaine CRD isolé, conduit cependant à la déstabilisation de la structure quaternaire de la Langérine. En effet, la région contenant la mutation est au coeur d'une interface impliquée dans le maintien de la forme trimérique de la Langérine en établissant des interactions avec le protomère voisin *via* notamment le résidu Phe241 et le résidu Lys200. Ainsi la forme mutante conduit majoritairement à un oligomère dimérique, cela explique la modification de la structure des BGs.

D'un point de vue cellulaire, les protéines mutantes sont correctement acheminées à la membrane cellulaire. Cette étude montre donc l'importance de la forme trimérique dans la formation de granule de Birbeck correctement structuré et met en évidence que la région de la mutation joue un rôle drastique dans le maintien de la structure quaternaire de la Langérine.

Enfin cette étude a permis de caractériser le domaine ECD de la Langérine en solution du point de vue structural. Le domaine neck apparaît plus flexible qu'imaginé précédemment au travers d'un modèle généré par notre équipe [100]. Sur l'ensemble de la séquence du domaine neck, certaines portions sont censées ne pas s'associer en coil-coiled et sont probablement à l'origine de points de flexibilité (Figure 9 D dans la publication du chapitre 11). L'organisation des têtes CRDs et le début du domaine neck de la structure trimérique, publiée par Feinberg et al, correspond bien à l'enveloppe SAXS que nous avons obtenue [38]. Cette flexibilité suggère une liberté de mouvement du domaine ECD exposé à la matrice extracellulaire pour reconnaître les agents pathogènes. Mais cette flexibilité, dans le cadre du granule de Birbeck structuré en CMS, ne semble pas cohérente à mon

---

sens avec des empilements de membranes réguliers et suggère une matrice de « rigidification ». Je reviendrai ultérieurement sur ce point pendant la discussion sur le nouveau site d'interaction indépendant du calcium de la Langérine pour les glycosaminoglycanes identifié dans ce travail de thèse.

### *Spécificité de la Langérine pour les sucres sulfatés*

Un deuxième axe majeur de mon travail de thèse fût de caractériser le mode d'interaction de la Langérine avec les glycosaminoglycanes. Cette étude se base sur la spécificité identifiée mais peu décrite, avant le début de ma thèse, pour la reconnaissance de sucres sulfatés [44]. Parmi les sucres sulfatés, nous nous sommes orientés vers les glycosaminoglycanes car ils sont abondants dans la matrice extracellulaire et exprimés par la majorité des cellules. De plus, un groupe de l'Institut s'intéresse aux différents rôles des glycosaminoglycanes dans la cellule et a donc développé des outils pour l'étude de ces sucres. Par le biais de cette collaboration, nous avons pu bénéficier de l'expertise de Romain Vivès.

Les premières expériences de SPR ont montré la capacité de reconnaissance de la Langérine pour l'héparine immobilisée sur une sensor chip. Nous avons démontré que cette interaction n'est pas dépendante du site calcique de la Langérine, car en présence de l'EDTA, nous observons aussi l'interaction (Figure2, manuscrit du chapitre 12).

Nous avons donc identifié un nouveau site de reconnaissance de la Langérine indépendant du calcium qui n'a jamais été encore décrit.

Ensuite, nous avons montré que cette interaction dépendait de l'oligomérisation de la protéine car aucune interaction n'a été observée dans le cas de la Langérine CRD, qui est monomérique (Figure2, manuscrit chapitre 12). Donc, la forme extracellulaire oligomérique de la Langérine est nécessaire pour l'interaction avec l'héparine. La forme trimérique semble donc être indispensable pour cette interaction. La caractérisation du degré d'oligomérisation de la Langérine nécessaire pour l'interaction avec l'héparine pourrait être entreprise, dans le futur, au travers d'expériences avec la Langérine possédant la mutation F241L, qui est majoritairement dimérique et fonctionnelle.

Pour évaluer la spécificité de ce nouveau site d'interaction, nous avons testé, par des expériences d'inhibition, les différents glycosaminoglycanes. En effet, les différents types de glycosaminoglycanes (héparane sulfate, chondroïtine sulfate ou kératane sulfate) possèdent un motif disaccharidique différent. Au sein d'un même type, la proportion des modifications (présence et position des groupements sulfates, épimérisation de l'acide glucuronique) présentes sur les sucres est différente. Une analyse de ces différentes proportions pour chaque type de GAGs, réalisée par nos collaborateurs, en complément des différences d'inhibition observées permet de suggérer des motifs favorisant cette interaction.

Tout d'abord, le motif IdoA2S-GlcNS6S est un motif important dans l'interaction

---

avec la Langérine et particulièrement le sulfate en position C6 de ce motif. En effet, le traitement de l'héparine par l'enzyme HSulf2, qui enlève le sulfate en position C6 du motif IdoA2S-GlcNS6S, conduit à la diminution de l'IC<sub>50</sub> de l'oligosaccharide par rapport à l'héparine non traitée.

Ensuite, nous n'avons pas observé de capacité d'interaction avec le kératane sulfate, dans la gamme testée, indiquant que cette structure saccharidique ne se lie pas au site d'interaction de l'héparine. Ceci n'est pas totalement étonnant car Tateno et al ont montré que le motif saccharidique présent dans le KS était reconnu dans le site calcique. Par contre, les glycosaminoglycanes de type chondroïtines sulfates sont reconnus dans le site caractérisé pour l'héparine car ils sont capables d'inhiber l'interaction de la Langérine avec l'héparine. Les chondroïtines sulfates sont intéressantes car elles possèdent différents motifs de sulfatation. Nous avons pu encore mettre en évidence que le sulfate en position C6 du galactosamine, cette fois-ci, est important pour l'interaction car la chondroïtine sulfate C est le meilleur inhibiteur compétitif de ce type de GAG et possède une proportion élevée en galactosamine sulfatée en C6. La comparaison entre la chondroïtine sulfate A (CS-A) et le dermatane sulfate (DS ou CS-B) montre l'importance de l'acide iduronique sulfaté en position C2. En effet, la différence principale entre ces deux chondroïtines sulfates est la présence plus importante d'acide iduronique sulfaté dans le DS.

Enfin, la différence majeure entre l'héparine et les chondroïtines sulfates, hormis le type des liaisons osidiques, est la présence d'un galactosamine au lieu du glucosamine pour les chondroïtines. Mais le premier sucre du motif répété de l'héparine et des chondroïtines sulfates est similaire et par conséquent l'acide glucuronique ou iduronique est aussi un maillon important dans ce mode d'interaction. Nous avons donc identifié dans ce travail un nouveau site d'interaction de la Langérine pour les GAGs de type héparane sulfate et chondroïtine sulfate. Les motifs fortement sulfatés le long de ces GAGs participent de façon importante à ces interactions et notamment le groupement sulfate en position C6 de la glucosamine ou de la galactosamine. Le seul facteur « limitant » pour l'interaction est la composition du disaccharide répété qui doit être composé d'un acide glucuronique ou iduronique lié à une glucosamine ou galactosamine sulfatée en position C6. La liaison osidique du motif disaccharidique des deux types de glycosaminoglycanes est différente et n'est donc pas un élément déterminant pour l'interaction avec les glycosaminoglycanes. D'ailleurs, il est probable que dans le site de fixation de ces GAGs, plusieurs modes d'interaction existent selon le GAG considéré.

Une étude de modélisation nous permet de proposer un site d'interaction de la Langérine avec l'héparine qui possède toutes les caractéristiques précédemment décrites (Figure 14.4). C'est-à-dire que le site ne fait pas intervenir le calcium dans l'interaction, et qu'il est nécessaire d'avoir une forme oligomérique de la protéine pour obtenir le site d'interaction.

---

Ainsi, pour conforter cette proposition, une étude de RMN STD avec un hexasaccharide de la série héparine est testé et montre une interaction non dépendante du calcium. L'étude du transfert de saturation des protons sur l'hexasaccharide et des calculs d'estimation de transfert de saturation à partir du modèle proposé peuvent nous permettre de conforter ou non la localisation du site. Outre la RMN, une approche cristallographique est à envisager à l'avenir avec un tel oligosaccharide. En effet, au cours de cette thèse, un groupe concurrent a montré la possibilité de cristalliser une forme tronquée et trimérique du domaine ECD de la Langérine [38]. Cette structure trimérique possède le site proposé par la modélisation et donc une approche de co-cristallisation devrait être tentée pour évaluer la pertinence de ce site.

Des expériences préliminaires de microscopie électronique et de SPR ont permis d'observer des propriétés d'assemblage de la Langérine. En effet, l'expérience de microscopie électronique suggère une capacité de formation de « chapelet » de la Langérine par l'héparine. L'expérience de SPR montre d'une part la capacité de la Langérine à interagir avec elle-même, et d'autre part que l'héparine est capable de stabiliser l'ensemble. Ces expériences sont à approfondir. Mais, ces résultats préliminaires suggèrent que la Langérine est capable à elle seule de s'assembler pour former le granule de Birbeck mais n'excluent pas le recours à une matrice saccharidique pour l'obtention d'une telle structure. L'interaction de la Langérine avec les glycosaminoglycanes dans le site proposé est un bon candidat dans l'idée d'une matrice de « rigidification » du domaine neck de la Langérine afin d'obtenir des empilements de membranes réguliers dans le granule de Birbeck. La validation du rôle des glycosaminoglycanes dans la formation et la structure du granule de Birbeck est envisageable par une étude de biologie cellulaire, couplée à la microscopie électronique, avec des cellules déficientes ou non en synthèse de GAGs. De telles expériences sont déjà envisagées pour la suite de ce travail.

De plus, nous avons montré dans l'étude précédente le rôle primordial de la forme trimérique de la Langérine dans la structure des BGs et dans le mode d'interaction avec les GAGs. À partir du modèle de l'organisation de la Langérine dans les BGs et de la modélisation avec l'héparine, l'hypothèse d'un rôle structurant des GAGs dans les BGs est théoriquement possible. En effet, en réalisant un modèle préliminaire, sans aucun calcul, on observe ainsi que les complexes macromoléculaires, impliquant des complexes GAG/Langérine (Figure 14.1 et 14.2), sont compatibles avec les données envisagées pour l'assemblage des granules de Birbeck (voir Figure 2.11 panneau C dans l'introduction).



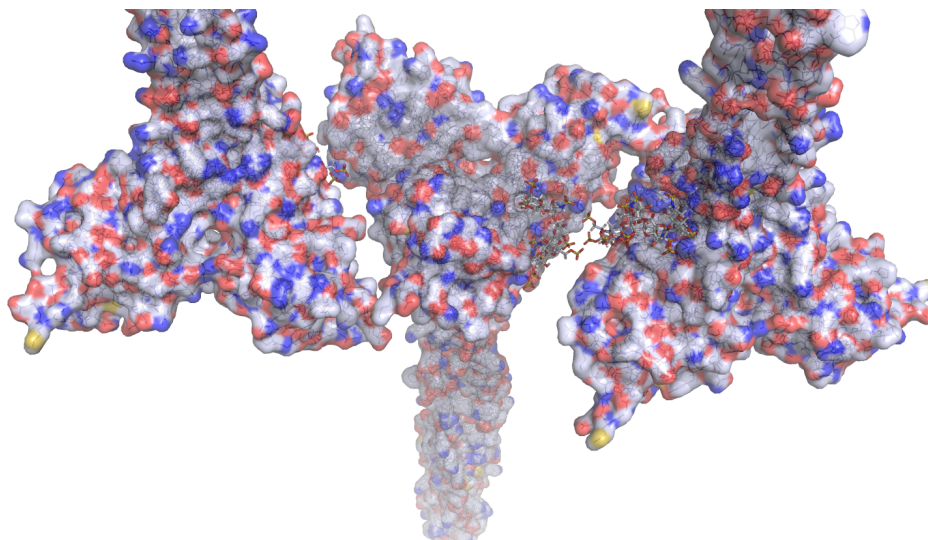


Figure 14.1 – Combinaison du modèle de l'organisation de la Langérine en alternance dans les BGs et de la modélisation avec l'héparine.

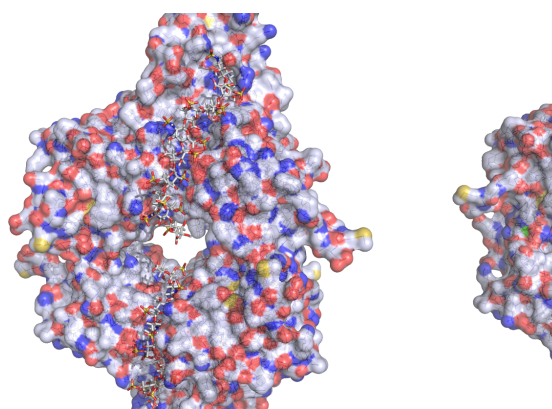


Figure 14.2 – Combinaison du modèle de l'organisation de la Langérine tête bêche dans les BGs et de la modélisation avec l'héparine.

Il a été montré que les glucosamines sulfatés de la chaîne d'héparine modulaient la migration des cellules de Langerhans [79] et certaines lectines sont caractérisées pour leur propriétés d'adhésion cellulaire ou de « rolling » [71] comme par exemple DC-SIGN. Il est donc aisé de penser que la Langérine peut être impliquée dans ce mécanisme de migration.

Enfin, les protéoglycanes et la Langérine sont fortement présents à la surface des cellules de Langerhans. Par conséquent, si la capacité d'interaction de la Langérine avec les GAGs est vérifiée *in cellulo*, alors la probabilité d'un partenaire protéoglycanique est grande dans le contexte physiologique dans lequel évolue les cellules de Langerhans. L'identification d'un tel partenaire permettrait peut être de mieux comprendre les différences dans la signalisation cellulaire de la Langérine dans les différents types de cellules

---

dendritiques. De plus, certaines études ont montré l'importance des héparanes sulfates à la surface des cellules dendritiques dermiques dans la reconnaissance de vésicules ressemblant au papillomavirus (« papillomavirus-like particle ») [25] et une co-localisation de ces vésicules à la surface des cellules de Langerhans [118]. Les perspectives dans la compréhension des voies de signalisation induites par la Langérine sont nombreuses si l'interaction entre la Langérine et les GAGs est avérée *in cellulo*.

### ***Interaction avec les glucosamines sulfatés dans le site calcique de la Langérine***

Un dernier axe de travail a été abordé pendant mon travail de thèse pour caractériser la spécificité du site calcique de la Langérine pour des glucosamines sulfatées. Tout d'abord, nous avons confirmé la capacité de reconnaissance de la Langérine pour des glucosamines sulfatées comme montré en 2010 par Tateno et al [99]. Ensuite, nous avons pu montrer que le groupement sulfate en C6 permet d'améliorer significativement l'affinité par rapport au GlcNac. Ceci montre une spécificité forte de la Langérine pour les glucosamines sulfatées en C6. Comme dans le cas du galactose sulfaté, précédemment publié, nous avons montré que le résidu de lysine 313 permet d'obtenir cette spécificité. En effet, la comparaison des structures cristallines de la Langérine CRD en présence de GlcNac et GlcNac6S nous a permis de mettre en évidence l'importance de groupement sulfate pour imposer une rotation de 180° du sucre afin d'établir une liaison électrostatique forte avec le résidu de lysine 313. Ce résidu de lysine n'est pas conservé dans les autres lectines de type-C. Sa présence permet probablement d'obtenir une spécificité unique de la Langérine pour les galactoses sulfatés en C6 et les glucosamines sulfatés en C6 dans le site calcique..

Dans le cas d'oligosaccharide, l'orientation du glucosamine sulfaté contraint d'effectuer une liaison osidique en bêta de l'hydroxyle anomérique de la glucosamine. Dans le cas des trisaccharides étudiés, cette liaison osidique est en alpha et par conséquent ceci nécessite une orientation du sucre à l'extrémité non réductrice similaire à celle du GlcNac. Ceci conduit à l'orientation de deux autres monosaccharides au niveau du solvant et peut expliquer l'affinité similaire du Tri1a et du GlcNac. Cette faible affinité est suprenante car par RMN, nous avons observé que tout le long du Tri1a, le transfert de saturation est quantifié autour de 50% et par conséquent indique que le sucre n'est pas en repoussé par la protéine et est suffisamment proche pour un transfert de saturation. Par cette même logique, les informations obtenues sur le Tri2b par RMN suggèrent que l'affinité peut être moins bonne car il y a moins de transfert de saturation le long du trisaccharide que pour le Tri1a. Par contre, ce phénomène peut être compensé par une « meilleure » interaction au niveau du sucre à l'extrémité non réductrice. En effet, le transfert de saturation est plus important au niveau notamment des hydrogènes du carbone 6 du sucre à l'extrémité non réductrice du Tri2b par rapport au Tri1a.

---

Dans le contexte physiologique, il reste à identifier des ligands pour émettre des hypothèses. Ainsi, des structures saccharidiques possédant des galactoses ou des glucosamines sulfatées ont été retrouvées au niveau de glycoprotéines d'enveloppe de virus comme par exemple pour la glycoprotéine d'enveloppe du virus de la grippe [60]. La caractérisation de structures saccharidiques sulfatées à la surface de virus permettrait d'identifier des ligands potentiels de la Langérine et ainsi approfondir les connaissances sur les agents pathogènes potentiellement reconnus.

Enfin, des motifs lewis<sup>X</sup> sulfatés et des glucosamines terminaux ont été décrits comme des marqueurs de certaines cellules cancéreuses [10, 86]. Peut-on penser que les cellules de Langerhans et la Langérine seraient impliqués dans des mécanismes de reconnaissance de ces cellules cancéreuses ? Peut-on imaginer aussi la possibilité d'utiliser la Langérine pour réaliser des tests de dépistage sur des biopsies ?

### *Perspectives en vue de développement d'inhibiteurs spécifiques de DC-SIGN et non de la Langérine*

Cette étude de la spécificité d'interaction de la Langérine pour les sucres sulfatés dans le site calcique peut à l'inverse permettre de proposer une nouvelle stratégie pour le développement d'inhibiteurs spécifiques de DC-SIGN. En effet, le remplacement du groupement sulfate en C6 par des groupements chargés positivement permettrait probablement de générer des sucres ayant un handicap électrostatique empêchant la reconnaissance dans le site de la Langérine sans pour autant affecter la reconnaissance dans le site de DC-SIGN. Des stratégies dans ce sens sont développées au laboratoire.

Dans les étapes préliminaires de criblage de ligands spécifiques de DC-SIGN par rapport à la Langérine, deux séries de composés à base « fucose » ou « mannose » ont été testées au laboratoire. Par le biais de la production de Langérine, j'ai pu contribuer modestement à permettre la réalisation de ces tests pour ces composés vis-à-vis de la Langérine. Ces travaux ont conduit à identifier des glycomimétiques d'intérêt dont l'optimisation va être poursuivie au laboratoire. Enfin, la stratégie abordée par ce réseau européen s'avère être une perspective dans la prévention contre les infections par le VIH. En effet, la deuxième étude relatée dans ce mémoire démontre la capacité d'inhiber la transmission du virus dans les muqueuses génitales par ces « pseudo-trimannoses ».

Par cette contribution, j'ai été associé aux publications respectives de ces travaux.

---

Sixième partie

Bibliographie

---

# Bibliographie

- [1] Jean-François Arrighi, Marjorie Pion, Eduardo Garcia, Jean-Michel Escola, Yvette van Kooyk, Teunis B Geijtenbeek, and Vincent Piguet. DC-SIGN-mediated infectious synapse formation enhances X4 HIV-1 transmission from dendritic cells to T cells. *The Journal of experimental medicine*, 200(10) :1279–1288, November 2004.
- [2] Sarolta Bacsa, Ghadah Karasneh, Sandor Dosa, Jian Liu, Tibor Valyi-Nagy, and Deepak Shukla. Syndecan-1 and syndecan-2 play key roles in herpes simplex virus type-1 infection. *The Journal of general virology*, 92(Pt 4) :733–743, April 2011.
- [3] J Banchereau and R M Steinman. Dendritic cells and the control of immunity. *Nature*, 392(6673) :245–252, 1998.
- [4] J Banchereau, F Briere, C Caux, J Davoust, S Lebecque, Y J Liu, B Pulendran, and K Palucka. Immunobiology of dendritic cells. *Annual review of immunology*, 18 :767–811, 2000.
- [5] J Magarian Blander and Ruslan Medzhitov. Regulation of phagosome maturation by signals from toll-like receptors. *Science (New York, N.Y.)*, 304(5673) :1014–1018, May 2004.
- [6] Michael D Bobardt, Andrew C S Saphire, Hsiu-Cheng Hung, Xiaocong Yu, Bernadette Van der Schueren, Zhe Zhang, Guido David, and Philippe A Gally. Syndecan captures, protects, and transmits HIV to T lymphocytes. *Immunity*, 18(1) :27–39, January 2003.
- [7] Morgane Bomsel and Annette Alfsen. Entry of viruses through the epithelial barrier : pathogenic trickery. *Nature reviews. Molecular cell biology*, 4(1) :57–68, January 2003.
- [8] A S BREATHNACH. OBSERVATIONS ON CYTOPLASMIC ORGANELLES IN LANGERHANS CELLS OF HUMAN EPIDERMIS. *Journal of anatomy*, 98 :265–270, April 1964.

- [9] Hannes E Bülow and Oliver Hobert. The molecular diversity of glycosaminoglycans shapes animal development. *Annual review of cell and developmental biology*, 22 : 375–407, 2006.
- [10] C Capon, J M Wieruszeski, J Lemoine, J C Byrd, H Leffler, and Y S Kim. Sulfated lewis X determinants as a major structural motif in glycans from LS174T-HM7 human colon carcinoma mucin. *The Journal of biological chemistry*, 272(51) :31957–31968, December 1997.
- [11] B Casu and U Lindahl. Structure and biological interactions of heparin and heparan sulfate. *Advances in carbohydrate chemistry and biochemistry*, 57 :159–206, 2001.
- [12] Marielle Cavrois, Jason Neidleman, Jason F Kreisberg, and Warner C Greene. In vitro derived dendritic cells trans-infect CD4 T cells primarily with surface-bound HIV-1 virions. *PLoS pathogens*, 3(1) :e4, January 2007.
- [13] M Cella, D Jarrossay, F Facchetti, O Alebardi, H Nakajima, A Lanzavecchia, and M Colonna. Plasmacytoid monocytes migrate to inflamed lymph nodes and produce large amounts of type I interferon. *Nature medicine*, 5(8) :919–923, 1999.
- [14] Lorenz Chatwell, Andrea Holla, Benedikt B Kaufer, and Arne Skerra. The carbohydrate recognition domain of Langerin reveals high structural similarity with the one of DC-SIGN but an additional, calcium-independent sugar-binding site. *Molecular immunology*, 45(7) :1981–1994, 2008.
- [15] V Claus, A Jahraus, T Tjelle, and T Berg. Lysosomal Enzyme Trafficking between Phagosomes, Endosomes, and Lysosomes in J774 Macrophages. *Journal of Biological ...*, 1998.
- [16] Collaborative Computational Project, Number 4. The CCP4 suite : programs for protein crystallography. *Acta crystallographica. Section D, Biological crystallography*, 50(Pt 5) :760–763, September 1994.
- [17] Matthew A Cooper. Optical biosensors in drug discovery. *Nature reviews. Drug discovery*, 1(7) :515–528, July 2002.
- [18] V David, F Hochstenbach, S Rajagopalan, and M B Brenner. Interaction with newly synthesized and retained proteins in the endoplasmic reticulum suggests a chaperone function for human integral membrane protein IP90 (calnexin). *The Journal of biological chemistry*, 268(13) :9585–9592, May 1993.



- [19] Bärbel I de Bakker, Frank de Lange, Alessandra Cambi, Jeroen P Korterik, Erik M H P van Dijk, Niek F van Hulst, Carl G Figdor, and Maria F Garcia-Parajo. Nanoscale organization of the pathogen receptor DC-SIGN mapped by single-molecule high-resolution fluorescence microscopy. *Chemphyschem : a European journal of chemical physics and physical chemistry*, 8(10) :1473–1480, July 2007.
- [20] Marein A W P de Jong, Lot de Witte, Menno J Oudhoff, Sonja I Gringhuis, Philippe Gallay, and Teunis B H Geijtenbeek. TNF-alpha and TLR agonists increase susceptibility to HIV-1 transmission by human Langerhans cells ex vivo. *The Journal of clinical investigation*, 118(10) :3440–3452, 2008.
- [21] Marein A W P de Jong, Lot de Witte, Saskia J A M Santegoets, Donna Fluitsma, Maureen E Taylor, Tanja D de Gruijl, and Teunis B H Geijtenbeek. Mutz-3-derived Langerhans cells are a model to study HIV-1 transmission and potential inhibitors. *Journal of leukocyte biology*, 87(4) :637–643, 2010.
- [22] Marein A W P de Jong, Lot de Witte, Maureen E Taylor, and Teunis B H Geijtenbeek. Herpes simplex virus type 2 enhances HIV-1 susceptibility by affecting Langerhans cell function. *Journal of immunology (Baltimore, Md : 1950)*, 185(3) :1633–1641, 2010.
- [23] Marein A W P de Jong, Lianne E M Vriend, Bart Theelen, Maureen E Taylor, Donna Fluitsma, Teun Boekhout, and Teunis B H Geijtenbeek. C-type lectin Langerin is a beta-glucan receptor on human Langerhans cells that recognizes opportunistic and pathogenic fungi. *Molecular immunology*, 47(6) :1216–1225, 2010.
- [24] Lot de Witte, Alexey Nabatov, Marjorie Pion, Donna Fluitsma, Marein A W P de Jong, Tanja de Gruijl, Vincent Piguët, Yvette van Kooyk, and Teunis B H Geijtenbeek. Langerin is a natural barrier to HIV-1 transmission by Langerhans cells. *Nature medicine*, 13(3) :367–371, 2007.
- [25] Lot de Witte, Younes Zoughlami, Birgit Aengeneyndt, Guido David, Yvette van Kooyk, Lutz Gissmann, and Teunis B H Geijtenbeek. Binding of human papilloma virus L1 virus-like particles to dendritic cells is mediated through heparan sulfates and induces immune activation. *Immunobiology*, 212(9-10) :679–691, 2007.
- [26] Lot de Witte, Alexey Nabatov, and Teunis B H Geijtenbeek. Distinct roles for DC-SIGN+-dendritic cells and Langerhans cells in HIV-1 transmission. *Trends in molecular medicine*, 14(1) :12–19, 2008.

- [27] Marc Delcommenne, Reiji Kannagi, and Pauline Johnson. TNF-alpha increases the carbohydrate sulfation of CD44 : induction of 6-sulfo N-acetyl lactosamine on N- and O-linked glycans. *Glycobiology*, 12(10) :613–622, October 2002.
- [28] Angeles Dominguez-Soto, Laura Aragoneses-Fenoll, Enrique Martin-Gayo, Lorena Martinez-Prats, Maria Colmenares, Marisa Naranjo-Gomez, Francesc E Borrás, Pilar Munoz, Mercedes Zubiaur, Maria L Toribio, Rafael Delgado, and Angel L Corbí. The DC-SIGN-related lectin LSECtin mediates antigen capture and pathogen binding by human myeloid cells. *Blood*, 109(12) :5337–5345, June 2007.
- [29] Sarah L Doyle and Luke A J O’Neill. Toll-like receptors : from the discovery of NFkappaB to new insights into transcriptional regulations in innate immunity. *Biochemical pharmacology*, 72(9) :1102–1113, October 2006.
- [30] Glenn Dranoff. Cytokines in cancer pathogenesis and cancer therapy. *Nature reviews. Cancer*, 4(1) :11–22, January 2004.
- [31] K Drickamer. C-type lectin-like domains. *Current opinion in structural biology*, 9(5) :585–590, October 1999.
- [32] J S Duke-Cohan, W Tang, and S F Schlossman. Attractin : a cub-family protease involved in T cell-monocyte/macrophage interactions. *Advances in experimental medicine and biology*, 477 :173–185, 2000.
- [33] Piotr Dziegiel, Barbara Dolilińska-Krajewska, Malgorzata Dumańska, Jadwiga Weclawek, Michał Jeleń, Marzena Podhorska-Okołów, Ewa Jagoda, Magdalena Fic, and Maciej Zabel. Coexpression of CD1a, langerin and Birbeck’s granules in Langerhans cell histiocytoses (LCH) in children : ultrastructural and immunocytochemical studies. *Folia histochemica et cytobiologica / Polish Academy of Sciences, Polish Histochemical and Cytochemical Society*, 45(1) :21–25, 2007.
- [34] A Dzionek, Y Sohma, J Nagafune, M Cella, M Colonna, F Facchetti, G Günther, I Johnston, A Lanzavecchia, T Nagasaka, T Okada, W Vermi, G Winkels, T Yamamoto, M Zysk, Y Yamaguchi, and J Schmitz. BDCA-2, a novel plasmacytoid dendritic cell-specific type II C-type lectin, mediates antigen capture and is a potent inhibitor of interferon alpha/beta induction. *The Journal of experimental medicine*, 194(12) :1823–1834, 2001.
- [35] Anneke Engering, Teunis B H Geijtenbeek, Sandra J van Vliet, Mietske Wijers, Ellis van Liempt, Nicolas Demaurex, Antonio Lanzavecchia, Jack Fransen, Carl G Figdor, Vincent Piguet, and Yvette van Kooyk. The dendritic cell-specific adhesion receptor

- DC-SIGN internalizes antigen for presentation to T cells. *Journal of immunology (Baltimore, Md : 1950)*, 168(5) :2118–2126, March 2002.
- [36] Emanuele Fanales-Belasio, Mariangela Raimondo, Barbara Suligoi, and Stefano Buttò. HIV virology and pathogenetic mechanisms of infection : a brief overview. *Annali dell'Istituto superiore di sanità*, 46(1) :5–14, 2010.
- [37] Hadar Feinberg, Maureen E Taylor, and William I Weis. Scavenger receptor C-type lectin binds to the leukocyte cell surface glycan Lewis(x) by a novel mechanism. *The Journal of biological chemistry*, 282(23) :17250–17258, June 2007.
- [38] Hadar Feinberg, Alex S Powlesland, Maureen E Taylor, and William I Weis. Trimeric structure of langerin. *The Journal of biological chemistry*, 285(17) :13285–13293, 2010.
- [39] Hadar Feinberg, Maureen E Taylor, Nahid Razi, Ryan McBride, Yuriy A Knirel, Sarah A Graham, Kurt Drickamer, and William I Weis. Structural basis for langerin recognition of diverse pathogen and mammalian glycans through a single binding site. *Journal of molecular biology*, 405(4) :1027–1039, 2011.
- [40] Hadar Feinberg, Maureen E Taylor, Nahid Razi, Ryan McBride, Yuriy A Knirel, Sarah A Graham, Kurt Drickamer, and William I Weis. Structural basis for langerin recognition of diverse pathogen and mammalian glycans through a single binding site. *Journal of molecular biology*, 405(4) :1027–1039, January 2011.
- [41] C G Figdor. Molecular characterization of dendritic cells operating at the interface of innate or acquired immunity. *Pathologie-biologie*, 51(2) :61–63, March 2003.
- [42] Vincent Flacher, Marielle Bouschbacher, Estelle Verronèse, Catherine Massacrier, Vanja Sisirak, Odile Berthier-Vergnes, Blandine de Saint-Vis, Christophe Caux, Collette Dezutter-Dambuyant, Serge Lebecque, and Jenny Valladeau. Human Langerhans cells express a specific TLR profile and differentially respond to viruses and Gram-positive bacteria. *Journal of immunology (Baltimore, Md : 1950)*, 177(11) : 7959–7967, 2006.
- [43] Vincent Flacher, Christoph H Tripp, Patrizia Stoitzner, Bernhard Haid, Susanne Ebner, Barbara Del Frari, Franz Koch, Chae Gyu Park, Ralph M Steinman, Juliana Idoyaga, and Nikolaus Romani. Epidermal Langerhans cells rapidly capture and present antigens from C-type lectin-targeting antibodies deposited in the dermis. *The Journal of investigative dermatology*, 130(3) :755–762, 2010.

- [44] Christine Galustian, Chae Gyu Park, Wengang Chai, Makato Kiso, Sandra A Brue-ning, Young-Sun Kang, Ralph M Steinman, and Ten Feizi. High and low affinity carbohydrate ligands revealed for murine SIGN-R1 by carbohydrate array and cell binding approaches, and differing specificities for SIGN-R3 and langerin. *International immunology*, 16(6) :853–866, June 2004.
- [45] Neha S Gandhi and Ricardo L Mancera. The structure of glycosaminoglycans and their interactions with proteins. *Chemical biology & drug design*, 72(6) :455–482, December 2008.
- [46] T B Geijtenbeek, D S Kwon, R Torensma, S J van Vliet, G C van Duijnhoven, J Middel, I L Cornelissen, H S Nottet, V N KewalRamani, D R Littman, C G Figdor, and Y van Kooyk. DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. *Cell*, 100(5) :587–597, March 2000.
- [47] Teunis B H Geijtenbeek, Gerard C F van Duijnhoven, Sandra J van Vliet, Elmar Krieger, Gert Vriend, Carl G Figdor, and Yvette van Kooyk. Identification of different binding sites in the dendritic cell-specific receptor DC-SIGN for intercellular adhesion molecule 3 and HIV-1. *The Journal of biological chemistry*, 277(13) : 11314–11320, March 2002.
- [48] Snigdha Ghosh, Claudia Bandulet, and Lars Nitschke. Regulation of B cell development and B cell signalling by CD22 and its ligands alpha2,6-linked sialic acids. *International immunology*, 18(4) :603–611, April 2006.
- [49] G Grouard, M C Rissoan, L Filgueira, I Durand, J Banchereau, and Y J Liu. The enigmatic plasmacytoid T cells develop into dendritic cells with interleukin (IL)-3 and CD40-ligand. *The Journal of experimental medicine*, 185(6) :1101–1111, 1997.
- [50] Karolin Hijazi, Yufei Wang, Carlo Scala, Simon Jeffs, Colin Longstaff, Daniel Stieh, Beth Haggarty, Guido Vanham, Dominique Schols, Jan Balzarini, Ian M Jones, James Hoxie, Robin Shattock, and Charles G Kelly. DC-SIGN increases the affinity of HIV-1 envelope glycoprotein interaction with CD4. *PloS one*, 6(12) :e28307, 2011.
- [51] Florian Hladik and M Juliana McElrath. Setting the stage : host invasion by HIV. *Nature reviews Immunology*, 8(6) :447–457, June 2008.
- [52] Ashleigh Hodges, Katherine Sharrocks, Mariola Edelmann, Dilair Baban, Arnaud Moris, Olivier Schwartz, Hal Drakesmith, Kay Davies, Benedikt Kessler, Andrew McMichael, and Alison Simmons. Activation of the lectin DC-SIGN induces an immature dendritic cell phenotype triggering Rho-GTPase activity required for HIV-1 replication. *Nature immunology*, 8(6) :569–577, June 2007.

- 
- [53] Andrea Holla and Arne Skerra. Comparative analysis reveals selective recognition of glycans by the dendritic cell receptors DC-SIGN and Langerin. *Protein engineering, design & selection : PEDS*, 2011.
- [54] Julie E Horbul, Stephen C Schmechel, Barrie R L Miller, Stephen A Rice, and Peter J Southern. Herpes simplex virus-induced epithelial damage and susceptibility to human immunodeficiency virus type 1 infection in human cervical organ culture. *PloS one*, 6(7) :e22638, 2011.
- [55] Robert E Hunger, Peter A Sieling, Maria Teresa Ochoa, Makoto Sugaya, Anne E Burdick, Thomas H Rea, Patrick J Brennan, John T Belisle, Andrew Blauvelt, Steven A Porcelli, and Robert L Modlin. Langerhans cells utilize CD1a and langerin to efficiently present nonpeptide antigens to T cells. *The Journal of clinical investigation*, 113(5) :701–708, March 2004.
- [56] Juliana Idoyaga, Cheolho Cheong, Koji Suda, Nao Suda, Jae Y Kim, Haekyung Lee, Chae Gyu Park, and Ralph M Steinman. Cutting edge : langerin/CD207 receptor on dendritic cells mediates efficient antigen presentation on MHC I and II products in vivo. *Journal of immunology (Baltimore, Md : 1950)*, 180(6) :3647–3650, 2008.
- [57] R L Jackson, S J Busch, and A D Cardin. Glycosaminoglycans : molecular properties, protein interactions, and role in physiological processes. *Physiological reviews*, 71(2) :481–539, April 1991.
- [58] Wolfgang Kabsch. XDS. *Acta crystallographica. Section D, Biological crystallography*, 66(Pt 2) :125–132, February 2010.
- [59] Daniel H Kaplan. In vivo function of Langerhans cells and dermal dendritic cells. *Trends in immunology*, 31(12) :446–451, 2010.
- [60] V K Karaivanova and R G Spiro. Sulphation of N-linked oligosaccharides of vesicular stomatitis and influenza virus envelope glycoproteins : host cell specificity, subcellular localization and identification of substituted saccharides. *The Biochemical journal*, 329 ( Pt 3) :511–518, February 1998.
- [61] Tatsuyoshi Kawamura, Forrest O Gulden, Makoto Sugaya, David T McNamara, Debra L Borris, Michael M Lederman, Jan M Orenstein, Peter A Zimmerman, and Andrew Blauvelt. R5 HIV productively infects Langerhans cells, and infection levels are regulated by compound CCR5 polymorphisms. *Proceedings of the National Academy of Sciences of the United States of America*, 100(14) :8401–8406, July 2003.

- [62] Adrien Kissenpfennig, Smina Ait-Yahia, Valérie Clair-Moninot, Hella Stössel, Edgar Badell, Yann Bordat, Joanne L Pooley, Thierry Lang, Eric Prina, Isabelle Coste, Olivia Gresser, Toufic Renno, Nathalie Winter, Geneviève Milon, Ken Shortman, Nikolaus Romani, Serge Lebecque, Bernard Malissen, Sem Saeland, and Patrice Douillard. Disruption of the langerin/CD207 gene abolishes Birbeck granules without a marked loss of Langerhans cell function. *Molecular and cellular biology*, 25 (1) :88–99, January 2005.
- [63] Johanneke Kleinnijenhuis, Marije Oosting, Leo A B Joosten, Mihai G Netea, and Reinout Van Crevel. Innate immune recognition of Mycobacterium tuberculosis. *Clinical & developmental immunology*, 2011 :405310, 2011.
- [64] A R Kolatkar and W I Weis. Structural basis of galactose recognition by C-type animal lectins. *The Journal of biological chemistry*, 271(12) :6679–6685, March 1996.
- [65] Paul Langerhans. Ueber die Nerven der menschlichen Haut. *Archiv für Pathologische Anatomie und Physiologie und für Klinische Medizin*, 44(2-3) :325–337, 1868.
- [66] Jong-Sil Lee, Gyung Hyuck Ko, Ho Cheol Kim, In Seok Jang, Kyung-Nyeo Jeon, and Jeong-Hee Lee. Langerhans cell sarcoma arising from Langerhans cell histiocytosis : a case report. *Journal of Korean medical science*, 21(3) :577–580, 2006.
- [67] B Lemaitre, E Nicolas, L Michaut, J M Reichhart, and J A Hoffmann. The dorso-ventral regulatory gene cassette *spätzle*/Toll/cactus controls the potent antifungal response in *Drosophila* adults. *Cell*, 86(6) :973–983, September 1996.
- [68] J Longley, J Kraus, M Alonso, and R Edelson. Molecular cloning of CD1a (T6), a human epidermal dendritic cell marker related to class I MHC molecules. *The Journal of investigative dermatology*, 92(4) :628–631, April 1989.
- [69] A Lupas. Coiled coils : new structures and new functions. *Trends in biochemical sciences*, 21(10) :375–382, October 1996.
- [70] Airlie J McCoy, Ralf W Grosse-Kunstleve, Paul D Adams, Martyn D Winn, Laurent C Storoni, and Randy J Read. Phaser crystallographic software. *Journal of applied crystallography*, 40(Pt 4) :658–674, August 2007.
- [71] Rodger P McEver and Cheng Zhu. Rolling cell adhesion. *Annual review of cell and developmental biology*, 26 :363–396, November 2010.
- [72] Gérard Morel. *Visualization of receptors*. methods in light and electron microscopy. CRC, 1997.

- [73] M Moulard, H Lortat-Jacob, I Mondor, G Roca, R Wyatt, J Sodroski, L Zhao, W Olson, P D Kwong, and Q J Sattentau. Selective interactions of polyanions with basic surfaces on human immunodeficiency virus type 1 gp120. *Journal of virology*, 74(4) :1948–1960, February 2000.
- [74] Masahiro Nakayama, Kuniaki Takahashi, Mitsuo Hori, Toshiyuki Okumura, Misutaki Saito, Mitsunori Yamakawa, Keiji Tabuchi, and Akira Hara. Langerhans cell sarcoma of the cervical lymph node : a case report and literature review. *Auris, nasus, larynx*, 37(6) :750–753, 2010.
- [75] Madelon Noordegraaf, Vincent Flacher, Patrizia Stoitzner, and Björn E Clausen. Functional redundancy of Langerhans cells and Langerin+ dermal dendritic cells in contact hypersensitivity. *The Journal of investigative dermatology*, 130(12) :2752–2759, 2010.
- [76] Youichi Ogawa, Tatsuyoshi Kawamura, Tetsuya Kimura, Masahiko Ito, Andrew Blauvelt, and Shinji Shimada. Gram-positive bacteria enhance HIV-1 susceptibility in Langerhans cells, but not in dendritic cells, via Toll-like receptor activation. *Blood*, 113(21) :5157–5166, 2009.
- [77] Caspar Ohnmacht, Andrea Pullner, Susan B S King, Ingo Drexler, Stefanie Meier, Thomas Brocker, and David Voehringer. Constitutive ablation of dendritic cells breaks self-tolerance of CD4 T cells and results in spontaneous fatal autoimmunity. *The Journal of experimental medicine*, 206(3) :549–559, 2009.
- [78] E K O’Shea, J D Klemm, P S Kim, and T Alber. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science (New York, N.Y.)*, 254(5031) :539–544, October 1991.
- [79] G M O’Sullivan, C M Boswell, and G M Halliday. Langerhans cell migration is modulated by N-sulfated glucosamine moieties in heparin. *Experimental dermatology*, 9(1) :25–33, February 2000.
- [80] A Poltorak, X He, I Smirnova, M Y Liu, C Van Huffel, X Du, D Birdwell, E Alejos, M Silva, C Galanos, M Freudenberg, P Ricciardi-Castagnoli, B Layton, and B Beutler. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice : mutations in Tlr4 gene. *Science (New York, N.Y.)*, 282(5396) :2085–2088, December 1998.
- [81] Guillaume Pompidor, Olivier Maury, Jean Vicat, and Richard Kahn. A dipicolinate lanthanide complex for solving protein structures using anomalous diffraction. *Acta crystallographica. Section D, Biological crystallography*, 66(Pt 7) :762–769, July 2010.

- [82] Alex S Powlesland, Tanja Fisch, Maureen E Taylor, David F Smith, Bérangère Tissot, Anne Dell, Stefan Pöhlmann, and Kurt Drickamer. A novel mechanism for LSEctin binding to Ebola virus surface glycoprotein through truncated glycans. *The Journal of biological chemistry*, 283(1) :593–602, January 2008.
- [83] Ehrhardt Proksch, Johanna M Brandner, and Jens-Michael Jensen. The skin : an indispensable barrier. *Experimental dermatology*, 17(12) :1063–1072, December 2008.
- [84] R W Sagebiel and T H Reed. Serial reconstruction of the characteristic granule of the Langerhans cell. *The Journal of cell biology*, 36(3) :595–602, March 1968.
- [85] Shinobu Saijo and Yoichiro Iwakura. Dectin-1 and Dectin-2 in innate immunity against fungi. *International immunology*, 23(8) :467–472, August 2011.
- [86] Tero Satomaa, Annamari Heiskanen, Iréne Leonardsson, Jonas Angström, Anne Olonen, Maria Blomqvist, Noora Salovuori, Caj Haglund, Susann Teneberg, Jari Natunen, Olli Carpén, and Juhani Saarinen. Analysis of the human cancer glycome identifies a novel group of tumor-associated N-acetylglucosamine glycan antigens. *Cancer research*, 69(14) :5811–5819, July 2009.
- [87] A Savina. Phagocytosis and antigen presentation in dendritic cells - Savina - 2007 - Immunological Reviews - Wiley Online Library. *Immunological reviews*, 2007.
- [88] Knut Schäkel, Reiji Kannagi, Bernhard Kniep, Yoshiko Goto, Chikako Mitsuoka, Jörg Zwirner, Afsaneh Soruri, Matthias von Kietzell, and E Rieber. 6-Sulfo LacNAc, a novel carbohydrate modification of PSGL-1, defines an inflammatory type of human dendritic cells. *Immunity*, 17(3) :289–301, September 2002.
- [89] Gerty Schreiber, Jurjen Tel, Kwinten H E W J Sliepen, Daniel Benitez-Ribas, Carl G Figdor, Gosse J Adema, and I Jolanda M de Vries. Toll-like receptor expression and function in human dendritic cell subsets : implications for dendritic cell-based anti-cancer immunotherapy. *Cancer immunology, immunotherapy : CII*, 59(10) :1573–1582, October 2010.
- [90] Meimei Shan, Per Johan Klasse, Kaustuv Banerjee, Antu K Dey, Sai Prasad N Iyer, Robert Dionisio, Dustin Charles, Lila Campbell-Gardener, William C Olson, Rogier W Sanders, and John P Moore. HIV-1 gp120 mannoses induce immunosuppressive responses from dendritic cells. *PLoS pathogens*, 3(11) :e169, November 2007.



- 
- [91] F P Siegal, N Kadowaki, M Shodell, P A Fitzgerald-Bocarsly, K Shah, S Ho, S Antonenko, and Y J Liu. The nature of the principal type 1 interferon-producing cells in human blood. *Science (New York, N. Y.)*, 284(5421) :1835–1837, 1999.
- [92] Nicola S Stambach and Maureen E Taylor. Characterization of carbohydrate recognition by langerin, a C-type lectin of Langerhans cells. *Glycobiology*, 13(5) :401–410, 2003.
- [93] R M Steinman. The dendritic cell system and its role in immunogenicity. *Annual review of immunology*, 9 :271–296, 1991.
- [94] R M Steinman and Z A Cohn. Identification of a novel cell type in peripheral lymphoid organs of mice. I. Morphology, quantitation, tissue distribution. *The Journal of experimental medicine*, 137(5) :1142–1162, May 1973.
- [95] R M Steinman and M D Witmer. Lymphoid dendritic cells are potent stimulators of the primary mixed leukocyte reaction in mice. *Proceedings of the National Academy of Sciences of the United States of America*, 75(10) :5132–5136, October 1978.
- [96] J Stetefeld, M Jenny, T Schulthess, R Landwehr, J Engel, and R A Kammerer. Crystal structure of a naturally occurring parallel right-handed coiled coil tetramer. *Nature structural & molecular biology*, 7(9) :772–776, September 2000.
- [97] Georges Tabarani, Michel Thépaut, David Stroebel, Christine Ebel, Corinne Vivès, Patrice Vachette, Dominique Durand, and Franck Fieschi. DC-SIGN neck domain is a pH-sensor controlling oligomerization : SAXS and hydrodynamic studies of extracellular domain. *The Journal of biological chemistry*, 284(32) :21229–21240, August 2009.
- [98] Li Tang, Juntao Yang, Xiaoming Tang, Wantao Ying, Xiaohong Qian, and Fuchu He. The DC-SIGN family member LSECTin is a novel ligand of CD44 on activated T cells. *European journal of immunology*, 40(4) :1185–1191, April 2010.
- [99] Hiroaki Tateno, Koji Ohnishi, Rikio Yabe, Norihito Hayatsu, Takashi Sato, Motohiro Takeya, Hisashi Narimatsu, and Jun Hirabayashi. Dual specificity of Langerin to sulfated and mannosylated glycans via a single C-type carbohydrate recognition domain. *The Journal of biological chemistry*, 285(9) :6390–6400, February 2010.
- [100] Michel Thépaut, Jenny Valladeau, Alessandra Nurisso, Richard Kahn, Bertrand Arnou, Corinne Vivès, Sem Saeland, Christine Ebel, Carine Monnier, Colette Dezutter-Dambuyant, Anne Imberty, and Franck Fieschi. Structural studies of langerin and

- Birbeck granule : a macromolecular organization model. *Biochemistry*, 48(12) : 2684–2698, 2009.
- [101] Christine Trumpfheller, Chae Gyu Park, Jennifer Finke, Ralph M Steinman, and Angela Granelli-Piperno. Cell type-dependent retention and transmission of HIV-1 by DC-SIGN. *International immunology*, 15(2) :289–298, February 2003.
- [102] Theodros Solomon Tsegaye and Stefan Pöhlmann. The multiple facets of HIV attachment to dendritic cell lectins. *Cellular microbiology*, 12(11) :1553–1561, November 2010.
- [103] Stuart G Turville, Paul U Cameron, Amanda Handley, George Lin, Stefan Pöhlmann, Robert W Doms, and Anthony L Cunningham. Diversity of receptors binding HIV on dendritic cell subsets. *Nature immunology*, 3(10) :975–983, October 2002.
- [104] J Valladeau, O Ravel, C Dezutter-Dambuyant, K Moore, M Kleijmeer, Y Liu, V Duvert-Frances, C Vincent, D Schmitt, J Davoust, C Caux, S Lebecque, and S Saeland. Langerin, a novel C-type lectin specific to Langerhans cells, is an endocytic receptor that induces the formation of Birbeck granules. *Immunity*, 12(1) : 71–81, 2000.
- [105] Jenny Valladeau, Valérie Clair-Moninot, Colette Dezutter-Dambuyant, Jean-Jacques Pin, Adrien Kissenpfennig, Marie-Genevieve Mattéi, Smina Ait-Yahia, Elizabeth E M Bates, Bernard Malissen, Franz Koch, François Fossiez, Nikolaus Romani, Serge Lebecque, and Sem Saeland. Identification of mouse langerin/CD207 in Langerhans cells and some dendritic cells of lymphoid tissues. *Journal of immunology (Baltimore, Md : 1950)*, 168(2) :782–792, 2002.
- [106] Jenny Valladeau, Colette Dezutter-Dambuyant, and Sem Saeland. Langerin/CD207 sheds light on formation of birbeck granules and their possible function in Langerhans cells. *Immunologic research*, 28(2) :93–107, 2003.
- [107] Pauline Verdijk, Remco Dijkman, Elsemieke I Plasmeijer, Aat A Mulder, Willem H Zoutman, A Mieke Mommaas, and Cornelis P Tensen. A lack of Birbeck granules in Langerhans cells is associated with a naturally occurring point mutation in the human Langerin gene. *The Journal of investigative dermatology*, 124(4) :714–717, April 2005.
- [108] Romain R Vivès, Rabia Sadir, Anne Imberty, Anna Rencurosi, and Hugues Lortat-Jacob. A kinetics and modeling study of RANTES(9-68) binding to heparin reveals

- a mechanism of cooperative oligomerization. *Biochemistry*, 41(50) :14779–14789, December 2002.
- [109] R Vollenweider and K Lennert. Plasmacytoid T-cell clusters in non-specific lymphadenitis. *Virchows Archiv. B, Cell pathology including molecular pathology*, 44(1) :1–14, 1983.
- [110] W I Weis and K Drickamer. Trimeric structure of a C-type mannose-binding protein. *Structure (London, England : 1993)*, 2(12) :1227–1240, December 1994.
- [111] W I Weis, M E Taylor, and K Drickamer. The C-type lectin superfamily in the immune system. *Immunological reviews*, 163 :19–34, June 1998.
- [112] Lin Weng, Roland Hübner, An Claessens, Patrick Smits, Jan Wauters, Przemko Tylzanowski, Eric Van Marck, and Jozef Merregaert. Isolation and characterization of chondrolectin (Chodl), a novel C-type lectin predominantly expressed in muscle cells. *Gene*, 308 :21–29, April 2003.
- [113] Scott C Wesselkamper, Bryan L Eppert, Gregory T Motz, Gee W Lau, Daniel J Hassett, and Michael T Borchers. NKG2D is critical for NK cell activation in host defense against *Pseudomonas aeruginosa* respiratory infection. *Journal of immunology (Baltimore, Md : 1950)*, 181(8) :5481–5489, October 2008.
- [114] K Wolff. The fine structure of the Langerhans cell granule. *The Journal of cell biology*, 35(2) :468–473, November 1967.
- [115] Li Wu and Vineet N KewalRamani. Dendritic-cell interactions with HIV : infection and viral dissemination. *Nature reviews Immunology*, 6(11) :859–868, 2006.
- [116] S J Wu, G Grouard-Vogel, W Sun, J R Mascola, E Brachtel, R Putvatana, M K Louder, L Filgueira, M A Marovich, H K Wong, A Blauvelt, G S Murphy, M L Robb, B L Innes, D L Birx, C G Hayes, and S S Frankel. Human skin Langerhans cells are targets of dengue virus infection. *Nature medicine*, 6(7) :816–820, July 2000.
- [117] S Yamada, Y Yamane, H Tsuda, K Yoshida, and K Sugahara. A major common trisulfated hexasaccharide core sequence, hexuronic acid(2-sulfate)-glucosamine(N-sulfate)-iduronic acid-N-acetylglucosamine-glucuronic acid-glucosamine(N-sulfate), isolated from the low sulfated irregular region of porcine intestinal heparin. *The Journal of biological chemistry*, 273(4) :1863–1871, January 1998.

- [118] Mengyong Yan, Judy Peng, Ibtissam A Jabbar, Xiaosong Liu, Luis Filgueira, Ian H Frazer, and Ranjeny Thomas. Despite differences between dendritic cells and Langerhans cells in the mechanism of papillomavirus-like particle antigen uptake, both cells cross-prime T cells. *Virology*, 324(2) :297–310, July 2004.
- [119] N E Zhou, C M Kay, and R S Hodges. Synthetic model proteins : the relative contribution of leucine residues at the nonequivalent positions of the 3-4 hydrophobic repeat to the stability of the two-stranded alpha-helical coiled-coil. *Biochemistry*, 31(25) :5739–5746, June 1992.
- [120] N E Zhou, B Y Zhu, C M Kay, and R S Hodges. The two-stranded alpha-helical coiled-coil is an ideal model for studying protein stability and subunit interactions. *Biopolymers*, 32(4) :419–426, April 1992.

## Septième partie

### Annexes

---

# Annexe 1

## Langérine S-CRD

Séquence nucléotidique :

ATGCAAAATGATATTCTACAGGTGGTTTCTCAAGGCTGGAAGTACTTCAAGGGGAAC TTC  
TATTACTTTTCTCTCATTCCAAAGACCTGGTATAGTGCCGAGCAGTTCTGTGTGTCCAGG  
AATTCACACCTGACCTCGGTGACCTCAGAGAGTGAGCAGGAGTTTCTGTATAAAACAGCG  
GGGGGACTCATCTACTGGATTGGCCTGACTAAAGCAGGGATGGAAGGGGACTGGTCCTGG  
GTGGATGACACGCCATTCAACAAGGTCCAAAGTGCGAGGTTCTGGATTCCAGGTGAGCCC  
AACAAATGCTGGGAACAATGAACACTGTGGCAATATAAAGGCTCCCTCACTTCAGGCCTGG  
AATGATGCCCCATGTGACAAAACGTTTCTTTTCATTTGTAAGCGACCCTATGTCCCATCA  
GAACCGTGA

Séquence protéique :

WSHPQFEKIEGRMNDILQVVSQGWKYFKGNFYFSLIPKTWYSAEQFCVSRNSHLTSVT  
SESEQEFLYKTAGGLIYWIGLTKAGMEGDWSWVDDTPFNKVQSARFWIPGEPNNAGNNEH  
CGNIKAPSLQAWNDAPCDKTFLFICKRPYPVSEP

La séquence en bleu correspond à la traduction de la séquence nucléotidique.

Les résidus en rouge correspondent aux mutations F241L et W264R

ProtParam :

Nombre d'acides aminés : 142.

Masse moléculaire : 16278,2.

pI théorique : 5,63.

Ala (A)	8	Cys (C)	4	His (H)	2	Met (M)	2	Thr (T)	7
Arg (R)	3	Gln (Q)	7	Ile (I)	7	Phe (F)	9	Trp (W)	7
Asn (N)	10	Glu (E)	8	Leu (L)	8	Pro (P)	9	Tyr (Y)	7
Asp (D)	6	Gly (G)	10	Lys (K)	9	Ser (S)	12	Val (V)	7

Nombre total d'acides aminés chargés négativement (Asp + Glu) : 14.

Nombre total d'acides aminés chargés positivement (Arg + Lys) : 12.

Dans le premier tableau toutes les Cys sont libres et dans le deuxième aucune ne l'est.

nm	276	278	279	280	282
$\epsilon$	48240	49254	49275	49030	47840
A à 1 mg / ml	2,963	3,026	3,027	3,012	2,939
nm	276	278	279	280	282
$\epsilon$	47950	49000	49035	48790	47600
A à 1 mg / ml	2,946	3,010	3,012	2,997	2,924



## Annexe 2

### Langérine ECD

Séquence nucléotidique :

ATGGGCACCATATCAGATGTAAAGACCAATGTCCAGTTGCTGAAAGGTCGTGTGGACAAC  
ATCAGCACCTGGATTCTGAAATTAAAAAGAATAGTGACGGCATGGAGGCAGCTGGCGTT  
CAGATCCAGATGGTGAATGAGAGCCTGGGTTATGTGCGTTCTCAGTTCCTGAAGTTAAAA  
ACCAGTGTGGAGAAGGCCAACGCACAGATCCAGATCTTAACAAGAAGTTGGGAAGAAGTC  
AGTACCTTAAATGCCCAAATCCCAGAGTTAAAAAGTGATTTGGAGAAAGCCAGTGCTTTA  
AATACAAAGATCCGGGCACTCCAGGGCAGCTTGGAGAATATGAGCAAGTTGCTCAAACGA  
CAAAATGATATTCTACAGGTGGTTTCTCAAGGCTGGAAGTACTTCAAGGGGAACCTCTAT  
TACTTTTCTCTCATTCCAAAGACCTGGTATAGTGCCGAGCAGTTCTGTGTGTCCAGGAAT  
TCACACCTGACCTCGGTGACCTCAGAGAGTGAGCAGGAGTTTCTGTATAAAACAGCGGGG  
GGACTCATCTACTGGATTGGCCTGACTAAAGCAGGGATGGAAGGGGACTGGTCCTGGGTG  
GATGACACGCCATTCAACAAGGTCCAAAGTGCGAGGTTCTGGATTCCAGGTGAGCCCAAC  
AATGCTGGGAACAATGAACACTGTGGCAATATAAAGGCTCCCTCACTTCAGGCCTGGAAT  
GATGCCCCATGTGACAAAACGTTTCTTTTCATTTGTAAGCGACCCTATGTCCCATCAGAA  
CCGTGA

Séquence protéique :

MGTISDVKTNVQLLKGRVDNISTLDSEIKKNSDGMEAAGVQIQMVNESLGYVRSQFLKLLK  
TSVEKANAQIQILTRSWEDEVSTLNAQIPELKSDLEKASALNTKIRALQGSLENMSKLLKR  
QNDILQVVSQGWKYFKGNFYFSLIPKTWYSAEQFCVSRNSHLTSVTSESEQEFYKTAG  
GLIYWIGLTKAGMEGDWSWVDDTPFNKVQSARFWIPGEPNNAGNNEHCGNIKAPSLQAWN  
DAPCDKTFLFICKRPYVPSEP\*

ProtParam :

Nombre d'acides aminés : 261.

Masse moléculaire : 29406,3.

pI théorique : 7,59.

Ala (A)	16	Cys (C)	4	His (H)	2	Met (M)	5	Thr (T)	14
Arg (R)	8	Gln (Q)	15	Ile (I)	15	Phe (F)	10	Trp (W)	8
Asn (N)	18	Glu (E)	17	Leu (L)	23	Pro (P)	10	Tyr (Y)	8
Asp (D)	11	Gly (G)	16	Lys (K)	21	Ser (S)	25	Val (V)	15

Nombre total d'acides aminés chargés négativement (Asp + Glu) : 28.

Nombre total d'acides aminés chargés positivement (Arg + Lys) : 29.

Dans le premier tableau toutes les Cys sont libres et dans le deuxième aucune ne l'est.

nm	276	278	279	280	282
ε	55090	56254	56280	56000	54640
A à 1 mg / ml	1,873	1,913	1,914	1,904	1,858
nm	276	278	279	280	282
ε	54800	56000	56040	55760	54400
A à 1 mg / ml	1,864	1,904	1,906	1,896	1,850

## Annexe 3

### Langérine S-ECD

Séquence nucléotidique :

ATGGGCACCATATCAGATGTAAAGACCAATGTCCAGTTGCTGAAAGGTCGTGTGGACAAC  
ATCAGCACCCCTGGATTCTGAAATTAAAAAGAATAGTGACGGCATGGAGGCAGCTGGCGTT  
CAGATCCAGATGGTGAATGAGAGCCTGGGTTATGTGCGTTCTCAGTTCCTGAAGTTAAAA  
ACCAGTGTGGAGAAGGCCAACGCACAGATCCAGATCTTAACAAGAAGTTGGGAAGAAGTC  
AGTACCTTAAATGCCCAAATCCCAGAGTTAAAAAGTGATTTGGAGAAAGCCAGTGCTTTA  
AATACAAAGATCCGGGCACTCCAGGGCAGCTTGGAGAATATGAGCAAGTTGCTCAAACGA  
CAAAATGATATTCTACAGGTGGTTTCTCAAGGCTGGAAGTACTTCAAGGGGAACTTCTAT  
TACTTTTCTCTCATTCCAAAGACCTGGTATAGTGCCGAGCAGTTCTGTGTGTCCAGGAAT  
TCACACCTGACCTCGGTGACCTCAGAGAGTGAGCAGGAGTTTCTGTATAAAACAGCGGGG  
GGACTCATCTACTGGATTGGCCTGACTAAAGCAGGGATGGAAGGGGACTGGTCCTGGGTG  
GATGACACGCCATTCAACAAGGTCCAAAGTGCGAGGTTCTGGATTCCAGGTGAGCCCAAC  
AATGCTGGGAACAATGAACACTGTGGCAATATAAAGGCTCCCTCACTTCAGGCCTGGAAT  
GATGCCCCATGTGACAAAACGTTTCTTTTCATTTGTAAGCGACCCTATGTCCCATCAGAA  
CCGTGA

Séquence protéique :

WSHPQFEKIEGRMGTISDVKTNVQLLKGRVDNISTLDSEIKKNSDGMEAAGVQIQMVN  
ESLGYVRSQFLKLKTSVEKANAQIQILTRSWEEVSTLNAQIPELKSDLEKASALNTKIRA  
LQGSLENMSKLLKRQNDILQVVSQGWKYFKGNFYFSLIPKTWYSAEQFCVSRNSHLT  
SVTSESEQEFLYKTAGGLIYWIGLTKAGMEGDWSWVD

La séquence en bleu correspond à la traduction de la séquence nucléotidique.

Protparam :

**Number of amino acids:** 273

**Molecular weight:** 30902.0

**Theoretical pI:** 7.78

**Amino acid composition:**

Ala (A)	16	5.9%
Arg (R)	9	3.3%
Asn (N)	18	6.6%
Asp (D)	11	4.0%
Cys (C)	4	1.5%
Gln (Q)	16	5.9%
Glu (E)	19	7.0%
Gly (G)	17	6.2%
His (H)	3	1.1%
Ile (I)	16	5.9%
Leu (L)	23	8.4%
Lys (K)	22	8.1%
Met (M)	5	1.8%
Phe (F)	11	4.0%
Pro (P)	11	4.0%
Ser (S)	26	9.5%
Thr (T)	14	5.1%
Trp (W)	9	3.3%
Tyr (Y)	8	2.9%
Val (V)	15	5.5%

**Total number of negatively charged residues (Asp + Glu):**  
30

**Total number of positively charged residues (Arg + Lys):**  
31

**Atomic composition:**

Carbon	C	1382
Hydrogen	H	2145
Nitrogen	N	371
Oxygen	O	416
Sulfur	S	9

**Formula:** C<sub>1382</sub>H<sub>2145</sub>N<sub>371</sub>O<sub>416</sub>S<sub>9</sub>

**Total number of atoms:** 4323

**Extinction coefficients:**

Extinction coefficients are in units of  $M^{-1} \text{ cm}^{-1}$ , at 280 nm measured in water.

Ext. coefficient      61670

Abs 0.1% (=1 g/l)    1.996, assuming ALL Cys residues  
appear as half cystines

Ext. coefficient      61420

Abs 0.1% (=1 g/l)    1.988, assuming NO Cys residues appear  
as half cystines

# Institut de Biologie Structurale

Laboratoire d'Enzymologie Moléculaire

Directrice : Nicole THIELENS



Grenoble 27 juillet 2009

to: Eric CHABROL  
IBS/LPM

From: Jean-Pierre ANDRIEU  
Laboratoire d'Enzymologie Moléculaire  
Institut de Biologie Structurale  
41 rue Jules Horowitz - 38027 Grenoble cedex 1

## Results of the N-terminal protein sequence analysis of sample:

**JP**

Sequence run #5767C

## Protein Sequence Determination

Amino acid sequence determination based on Edman degradation was performed using an Applied Biosystems gas-phase sequencer model 492 (s/n: 9510287J). Phenylthiohydantoin amino acid derivatives generated at each sequence cycle were identified and quantitated on-line with an Applied Biosystems Model 140C HPLC system using the data analysis system for protein sequencing from Applied Biosystems Model 610A (software version 2.1). The PTH-amino acid standard kit (Perkin-Elmer P/N 4340968) was used and reconstituted according to the manufacturer's instructions (900776 Rev D). The procedures and reagents used were as recommended by the manufacturer. Chromatography was used to identify and quantify the derivatized amino acid removed at each sequence cycle. Retention times and integration values of peaks were compared to the chromatographic profile obtained for a standard mixture of derivatized amino acids.

### Sequence run #5767C

**At least, four sequences were found:**

**V V S Q G W K Y**

**in accordance with the sequence of Lg-EC, from the residue 127,**

**D I L Q V V S Q**

**in accordance with the sequence of Lg-EC, from the residue 123,**

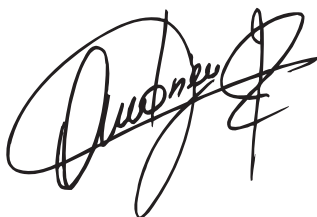
**S Q G W K Y F K**

**in accordance with the sequence of Lg-EC, from the residue 129, and**

**I L Q V V S Q G**

**in accordance with the sequence of Lg-EC, from the residue 124.**

Sequence determination analyses were performed by J.-P. Andrieu (Institut de Biologie Structurale, CEA/CNRS/UJF, Grenoble, France).

A handwritten signature in black ink, appearing to read 'Andrieu', with a stylized flourish at the end.

Jean-Pierre ANDRIEU

## **Articles :**

Andreini M, Doknic D, Sutkeviciute I, Reina JJ, Duan J, Chabrol E, Thepaut M, Moroni E, Doro F, Belvisi L, Weiser J, Rojo J, Fieschi F, Bernardi A. « Second generation of fucose-based DC-SIGN ligands: affinity improvement and specificity versus Langerin. » *Org Biomol Chem.* 2011 Aug 21;9(16):5778-86. Epub 2011 Jul 7.

Berzi A, Reina JJ, Ottria R, Sutkeviciute I, Antonazzo P, Sanchez-Navarro M, Chabrol E, Biasin M, Trabattoni D, Cetin I, Rojo J, Fieschi F, Bernardi A, Clerici M. « A glycomimetic compound inhibits DC-SIGN-mediated HIV infection in cellular and cervical explant models. » *AIDS.* 2012 Jan 14;26(2):127-37.

## **Articles en cours de publication :**

Eric Chabrol, Michel Thépaut, Colette Dezutter-Dambuyant, Corinne Vivès, Julien Marcoux, Richard Kahn, Jenny Valladeau-Guilemond, Patrice Vachette, Dominique Durand, and Franck Fieschi. « Alteration of langerin oligomerisation state affects gp120 recognition and Birbeck granules formation: effect of the F241L mutation. »

Eric Chabrol, Alessandra Nurisso, Antoine Daina, Emilie Vassal, Michel Thepaut, Eric Girard, Romain Vivès, Franck Fieschi. « A new  $\text{Ca}^{2+}$ -independent interacting mode of langerin with glycosaminoglycans contrast with gp120 binding mode. »

Michel Thépaut, Cinzia Guzzi, Ieva Sutkeviciute, Sara Sattin, Renato Ribeiro-Viana, Eric Chabrol, Javier Rojo, Jesus Angulo, Anna Bernardi, Pedro M. Nieto, and Franck Fieschi. « Crystal structure coupled to NMR analysis identifies the binding mode of pseudo-1,2-mannobioside in the carbohydrate recognition domain of the C-type lectin receptor DC-SIGN. »

## **Communications orales :**

Octobre 2008 : Communication orale au congrès Glucidoc.

Novembre 2010 : Communication orale à l'axe thématique « Immunité et Interactions Hôte-Pathogène » de l'Institut de Biologie Structurale.

Mai 2011 : Communication orale au congrès Glucidoc.

## **Poster :**

Juin 2009 : Poster à la Journée de l'Institut de Biologie Structurale (Grenoble)

Février 2010 : Poster à la journée du Partnership for Structural Biology (Grenoble)

Mai 2010 : Poster au Summer Course Glycosciences (Wageningen, Pays-Bas)

Mai 2011 : Poster au congrès Glucidoc (Sète, France)





## Résumé :

Les cellules dendritiques jouent un rôle primordial dans le système immunitaire. En effet, ces cellules sont à l'interface entre l'immunité innée et adaptative par leur capacité de reconnaissance, d'internalisation et de dégradation de pathogènes afin de présenter des antigènes aux lymphocytes. La capacité de reconnaissance est engendrée par l'expression de différents récepteurs à la surface de ces cellules. Parmi ces récepteurs, deux grandes familles permettent la reconnaissance d'un large panel de différents pathogènes, comme les TLRs (« Toll-Like Receptors ») et les lectines de type-C. Ces récepteurs sont utilisés comme marqueurs des différents sous-types de cellules dendritiques. Par exemple, parmi les lectines de type-C, DC-SIGN est majoritairement exprimée dans les cellules dendritiques dermiques alors que la Langerine est, quand à elle, fortement exprimée par les cellules dendritiques épidermiques, les cellules de Langerhans. Ces deux sous-types de cellules dendritiques divergent par leur réponse à l'infection par le VIH (« virus d'immunodéficience humaine »). En effet, le virus utilise DC-SIGN pour détourner le rôle de ces cellules afin d'infecter les lymphocytes T alors que la reconnaissance du VIH par la Langerine, dans les cellules de Langerhans, conduit à la clairance de virus par son internalisation dans le granule de Birbeck. Cet organite est spécifique des cellules de Langerhans et nécessite l'expression de la Langerine.

Ce travail de thèse s'est donc focalisé sur la caractérisation structurale et fonctionnelle de la Langerine. Il a permis de mettre en évidence l'importance de la structure tertiaire du domaine CRD et de la structure quaternaire de la protéine pour la formation et la bonne structuration du granule de Birbeck. Ensuite, l'étude fonctionnelle de cette lectine, notamment par résonance plasmonique de surface, nous a conduit à identifier une nouvelle spécificité de reconnaissance de la Langerine pour les glycosaminoglycanes dans un site d'interaction différent du site canonique. Enfin, nous avons caractérisé une spécificité de reconnaissance du site canonique pour les monosaccharides sulfatés de type glucosamine en utilisant la résonance plasmonique de surface et la cristallographie.

## Abstract :

Dendritic cells play a crucial role in the immune system. Indeed, these cells are at the interface between innate and acquired immunity by their capacities of recognition, internalisation and pathogen degradation to present antigens to T lymphocytes. The recognition capacity is generated by the expression of diverse receptors onto the cell surface. Among these receptors, two large families allow the recognition of a large panel of different pathogens, as TLRs ("Toll-Like Receptor") and C-type lectins. These receptors are used as markers of different dendritic cells subtypes. For example, and among the C-type lectins, DC-SIGN is mainly expressed onto dermic dendritic cells contrary to langerin, which is highly expressed onto epidermic dendritic cells, called Langerhans cells. These two subtypes of dendritic cells differ in their response of HIV infection. Indeed, the virus recognition by DC-SIGN enables hijacking the dendritic cell to infect T lymphocyte contrary to langerin recognition, in Langerhans cells, which allows the clearance of the virus by its internalisation into Birbeck granules. This organite is specific of Langerhans cells and requires langerin expression.

This work is focused on structural and functional characterisation of langerin. It highlights the importance of the CRD tertiary structure and the quaternary structure of the protein for the formation and the structure of Birbeck granules. Then, functional study by surface plasmon resonance enabled us to identify a new binding site of langerin for glycosaminoglycans. Finally, we have characterised a recognition specificity of langerin for sulphated monosaccharide of glucosamine type using surface plasmon resonance and crystallography.